



DISERTASI – SS143506

PENGEMBANGAN MODEL REGRESI NONPARAMETRIK BIRESPOK KONTINU MENGUNAKAN METODE MARS

Oleh :

Ayub Parlin Ampulembang
NRP. 1312301702

Promotor :

Dr. Bambang Widjanarko Otok, M.S

Co-Promotor :

Dr. Agnes Tuti Rumiati, M.Sc
Dr. Budiasih, SE, ME

PROGRAM DOKTOR
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2017

Halaman ini sengaja dikosongkan

LEMBAR PENGESAHAN

Disertasi disusun untuk memenuhi salah satu syarat memperoleh
gelar Doktor (Dr.)

di

Program Doktor Departemen Statistika
Institut Teknologi Sepuluh Nopember

Oleh:

Ayub Parlin Ampulembang
NRP. 1312301702

Tanggal Ujian : 3 Mei 2017

Periode Wisuda : September 2017

Disetujui oleh:

1. Dr. Bambang Widjanarko Otok, M.Si.
NIP. 19681124 199412 1 001

(Promotor)

2. Dr. Agnes Tuti Rumiati, M.Sc.
NIP. 19570724 198503 2 002

(Co-Promotor)

3. Dr. Budiasih, SE, ME.
NIP. 19610219 198312 2 001

(Co-Promotor)

4. Dr. Ir. Hari Wijayanto, M.Si
NIP. 19650421 199002 1 001

(Penguji
Eksternal)

5. Dr. Ir. Setiawan, MS.
NIP. 19601030 198701 1 001

(Penguji)

6. Prof. Dr. Drs. I Nyoman Budiantara, M.Si
NIP. 19650603 198903 1 003

(Penguji)

7. Dr. Suhartono, S.Si., M.Sc.
NIP. 19710929 199512 1 001

(Penguji)



Dekan Fakultas MIPA

Prof. Dr. Basuki Widodo, M.Sc.
NIP. 19650605 198903 1 002

Halaman ini sengaja dikosongkan

PENGEMBANGAN MODEL REGRESI NONPARAMETRIK BIRESPON KONTINU MENGGUNAKAN METODE MARS

Nama Mahasiswa : Ayub Parlin Ampulembang
NRP : 1312301702
Promotor : Dr. Bambang Widjanarko Otok, M.Si.
Co-Promotor : 1. Dr. Agnes Tuti Rumiati, M.Sc.
2. Dr. Budiasih, SE, ME.

ABSTRAK

Tujuan penelitian ini adalah mengembangkan model regresi nonparametrik birespon kontinu menggunakan metode MARS. Kajian awal difokuskan pada kajian teori tentang bentuk estimator MARS birespon kontinu. Hasil kajian teori selanjutnya diterapkan pada data simulasi dan data riil yaitu pada data tingkat kesejahteraan rakyat di Pulau Jawa. Hasil penelitian menunjukkan bahwa estimasi fungsi regresi MARS birespon kontinu diperoleh dengan metode *maximum likelihood estimation* (MLE). Untuk mengakomodir korelasi antar respon, maka estimasi fungsi regresi memuat matriks bobot berupa matriks kovarians *error*. Umumnya matriks bobot ini tidak diketahui, sehingga diestimasi dari data. Hasil empiris pada simulasi fungsi asli MARS menunjukkan bahwa semakin besar ukuran sampel (n) dan semakin kecil varians *error* maka akan menghasilkan estimasi parameter dan knot yang semakin dekat dengan nilai aktualnya. Metode ini cenderung baik digunakan untuk $n \geq 50$, dan kurang baik untuk $n < 50$ dengan varians *error* yang besar. Hasil simulasi juga menunjukkan bahwa estimator MARS birespon kontinu dapat mengestimasi berbagai bentuk fungsi percobaan, seperti fungsi Trigonometri, Polinomial, dan Eksponensial, yang ditunjukkan melalui pendekatan grafik. Hasil simulasi juga menunjukkan bahwa fungsi basis optimal dan korelasi yang tinggi pada model MARS birespon kontinu berperan penting dalam mendapatkan hasil estimasi yang lebih baik. Simulasi juga menunjukkan bahwa nilai mutlak korelasi diatas 0,3 ($|\rho| > 0,3$) cenderung memberikan hasil prediksi yang lebih baik. Hasil estimasi MARS birespon kontinu pada pemodelan data riil (pemodelan persentase penduduk miskin dan IPM) memberikan akurasi yang cukup baik dengan nilai R^2 sebesar 76,86% dan $RMSE$ sebesar 2,2221. Hasil validasi dengan nilai $RMSECV$ sebesar 2,7066 menunjukkan bahwa hasil estimasi cukup valid menggambarkan fenomena riil yang ada. Validasi juga menunjukkan bahwa model MARS pada persentase penduduk miskin dan IPM yang melibatkan korelasi antar respon menghasilkan tingkat prediksi lebih baik dibandingkan dengan tidak melibatkan korelasi.

Kata Kunci : MARS, Birespon, Bobot, Korelasi, MLE

DEVELOPMENT OF CONTINUE BIRESponses NONPARAMETRIC REGRESSION MODEL USING MARS

Name of Student : Ayub Parlin Ampulembang
Identity Number : 1312301702
Promotor : Dr. Bambang Widjanarko Otok, M.Si.
Co-Promotor : 1. Dr. Agnes Tuti Rumiati, M.Sc.
2. Dr. Budiasih, SE, ME.

ABSTRACT

The aim of this study is to develop a model of continue biresponses nonparametric regression using MARS. The first research is focused on the theory study to obtain estimator of continue biresponses MARS. Furthermore, the results of the theory study is applied to the simulation data and the real data i.e people welfare data in Java island. The results showed that the estimation of regression function in continue biresponses MARS is obtained by using maximum likelihood estimation (MLE) method. To accommodate the correlation between two responses, the estimation involves weighted matrix i.e. covarians error matrix. In general, this weighted matrix is unknown, so it must be estimated from the data. The simulation results on the original MARS function showed that the larger sample size (n) and the smaller variance error will result the closer estimation of parameters and knots respectively with its actual value. This method tends to be well used for $n \geq 50$ sample sizes, and tends to be poorly used for $n < 50$ with the large variance errors. The simulation results also showed that estimator of continue biresponses MARS can estimate various forms of the experiments function, such as in the experiment of trigonometry, polynomial and exponential function, which is shown by graph approach. Besides that, the simulation results also showed that the optimal basis functions and the higher absolute correlation play an important role in getting the better estimation in continue biresponses MARS model. Simulation also showed that absolute value of correlation above 0.3 ($|\rho| > 0.3$) tends to give better prediction result. Furthermore, the results on the real data i.e. the poverty rate and HDI modeling provides quite accurate model with R^2 values as 76.86% and RMSE value as 2.2221. The validation results with RMSECV value as 2.7066 indicates that the model estimation is quite valid to describe the existing real phenomenon. The validation results also showed that the modeling of poverty rate and IPM using MARS which involves its responses correlation produces better prediction compared to not involve correlation.

Keywords : MARS, Biresponses, Weighted, Correlation, MLE

KATA PENGANTAR

Segala puji dan syukur bagi Tuhan Yesus Kristus atas penyertaan, berkat dan kasih-Nya yang sempurna kepada penulis sehingga dapat menyelesaikan disertasi dengan judul **Pengembangan Model Regresi Nonparametrik Birespon Kontinu Menggunakan Metode MARS**.

Dalam penyusunan disertasi ini, seringkali ditemui berbagai kesulitan dan hambatan berkaitan dengan hal-hal baru yang ditemukan. Akan tetapi berkat bantuan dan bimbingan dari berbagai pihak, penulis dapat menyelesaikannya. Untuk itu penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Kepala BPS RI, Bapak Kepala BPS Provinsi Sulawesi Selatan dan Bapak Kepala BPS Kota Makassar Provinsi Sulawesi Selatan, yang telah memberikan ijin, kesempatan serta beasiswa kepada penulis untuk melanjutkan studi S3 di Departemen Statistika FMIPA ITS Surabaya.
2. Bapak Dr. Suhartono, S.Si., M.Sc., selaku Kepala Departemen Statistika FMIPA ITS Surabaya, yang telah memberikan ijin, bantuan dan dukungan fasilitas untuk kelancaran penulisan disertasi ini.
3. Bapak Dr. rer. pol. Heri Kuswanto, M.Si., selaku Kepala Program Studi Pascasarjana Statistika FMIPA ITS Surabaya, yang telah memberikan ijin, bantuan dan dukungan fasilitas untuk kelancaran penulisan disertasi ini.
4. Bapak Dr. Bambang Widjanarko Otok, M.Si., sebagai Promotor yang telah memberikan bimbingan, motivasi, arahan dan petunjuk yang sangat membantu dalam penulisan disertasi ini.
5. Ibu Dr. Agnes Tuti Rumiati, M.Sc. dan Ibu Dr. Budiasih, SE. ME., sebagai Co-Promotor yang telah memberikan bimbingan, motivasi, arahan dan petunjuk yang sangat membantu dalam penulisan disertasi ini.
6. Bapak Prof. Dr. Drs. I Nyoman Budiantara, M.Si., Bapak Dr. Ir. Setiawan, M.S. dan Bapak Dr. Suhartono, S.Si., M.Sc., sebagai tim penilai disertasi yang banyak memberikan masukan dan saran dalam perbaikan disertasi ini.
7. Bapak Dr. Ir. Hari Wijayanto, M.Si., sebagai penguji eksternal dari Institut Pertanian Bogor (IPB) yang telah memberikan saran dan masukan dalam penulisan disertasi ini.
8. Bapak Dr. Wahyu Wibowo, M.Si. dan Ibu Dr. Kartika Fithriasari, M.Si., sebagai tim validasi yang telah memberikan masukan dan saran dalam perbaikan disertasi ini.

9. Bapak Dr. Puhadi, M.Sc., sebagai dosen wali yang telah memberikan bimbingan kepada penulis selama menempuh studi S3 di ITS Surabaya.
10. Keluarga besar dosen Departemen Statistika FMIPA ITS Surabaya, yang telah memberikan bekal ilmu kepada penulis selama menempuh studi S3.
11. Keluarga besar staf tata usaha dan karyawan Departemen Statistika ITS Surabaya, yang memberikan kemudahan dan fasilitas selama studi S3.
12. Rekan-rekan mahasiswa Program Doktor Statistika ITS Surabaya, yang selalu memberikan dorongan dan bantuan pemikiran dalam penulisan disertasi ini.
13. Ibunda tercinta, D.S. Tandirerung yang banyak memberikan bantuan moril maupun materiil, doa, restu dan motivasi disaat penulis kehilangan harapan dan semangat, hingga disertasi ini dapat terselesaikan. Juga buat tante Milka Tandirerung yang senantiasa mendoakan terselesainya disertasi ini.
14. Ayah dan ibu mertua (L.R. Pararrungan dan D.S. Buli) yang senantiasa mendoakan dan memberikan motivasi dalam penyelesaian disertasi ini.
15. Teristimewa, istriku Asrihine Pararrungan dan anakku Aucky Ebenezer Ampulembang yang senantiasa menjadi sumber motivasi dan penyemangat untuk segera menyelesaikan disertasi ini.
16. Kakakku (Jusak Ampulembang dan Yenni Mangoting serta anak Rara dan Lola), kakakku (Uria Fatmariaty dan Yori A. Nawaraherja serta anak Bevan dan Felix), sahabatku Abraham Tuanakotta, serta semua keluarga yang tak henti-hentinya mendoakan dan memberi semangat dalam penyelesaian disertasi ini.
17. Terakhir, disertasi ini penulis dedikasikan kepada almarhum ayahanda tercinta, Lukas Ampulembang atas segala kasih sayang yang tak pernah bisa terbalaskan dan harapannya kepada penulis untuk terus melanjutkan pendidikan hingga setinggi-tingginya.

Akhirnya penulis berharap semoga disertasi ini dapat memberikan manfaat dan sumbangan bagi wawasan keilmuan. Amin.

Surabaya, 14 Juli 2017

Ayub Parlin Ampulembang

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	ii
ABSTRAK	iii
ABSTRACT	iv
KATA PENGANTAR	v
DAFTAR ISI	vii
DAFTAR GAMBAR	x
DAFTAR TABEL	xiv
DAFTAR LAMPIRAN	xv
DAFTAR LAMBANG DAN ARTI	xvi
BAB 1. PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Permasalahan	7
1.3. Tujuan Penelitian	8
1.4. Manfaat Penelitian	8
1.5. Orisinalitas Penelitian	9
1.6. Batasan Penelitian	11
BAB 2. KAJIAN PUSTAKA DAN DASAR TEORI	13
2.1. Analisis Regresi	13
2.1.1. Model Regresi Parametrik	13
2.1.2. Model Regresi Nonparametrik	14
2.2. <i>Truncated Spline</i> dalam Regresi Nonparametrik	15
2.3. <i>Multivariate Adaptive Regression Spline</i> (MARS)	21
2.3.1. Fungsi MARS	22
2.3.2. Estimasi Fungsi Regresi MARS pada Respon Kontinu	25
2.3.3. Pemilihan Fungsi Basis Optimal dalam MARS	27
2.4. Model Regresi Nonparametrik Birespon Kontinu	34
2.5. Ukuran Kesesuaian Model	36

2.6. Uji Linieritas Data dengan <i>Reset Test</i>	37
2.7. Uji Normal Multivariat Data dengan Q-Q Plot	38
2.8. Teorema dan Proposisi Matriks Aljabar	39
2.9. Definisi Operasional Variabel Penelitian	40
2.9.1. Persentase Penduduk Miskin (<i>Head Count Index</i>)	41
2.9.2. Indeks Pembangunan Manusia (IPM)	42
2.9.3. Pertumbuhan Ekonomi	44
2.9.4. Angka Beban Ketergantungan	45
2.9.5. Persentase Pengeluaran Perkapita untuk Makanan	47
2.9.6. Tingkat Pengangguran Terbuka (TPT)	48
2.9.7. Persentase Penduduk yang Pekerjaan Utamanya di Sektor Pertanian	49
BAB 3. METODOLOGI PENELITIAN	51
3.1. Estimasi Fungsi Regresi dalam Model MARS Birespon Kontinu	51
3.2. Estimasi Matriks Bobot	52
3.3. Penerapan Model MARS Birespon Kontinu pada Data Simulasi	53
3.4. Penerapan Model MARS Birespoon Kontinu pada Data Riil	56
3.4.1. Sumber Data	56
3.4.2. Variabel Penelitian	56
3.4.3. Struktur Data Penelitian	57
3.4.4. Langkah-langkah Analisis	57
BAB 4. HASIL KAJIAN TEORITIK	59
4.1. Estimasi Fungsi Regresi dalam Model MARS Birespon Kontinu	59
4.2. Estimasi Matriks Bobot	70
BAB 5. HASIL KAJIAN SIMULASI	79
5.1. Simulasi MARS Birespon Kontinu pada Fungsi Asli	79
5.1.1. Studi Simulasi pada Variasi Ukuran Sampel dan Varians <i>Error</i>	85
5.1.2. Studi Simulasi pada Variasi Ukuran Korelasi	95

5.2. Simulasi MARS Birespon Kontinu pada Berbagai Fungsi Percobaan	99
5.2.1. Simulasi MARS Birespon Kontinu pada Fungsi Trigonometri-Polinomial	100
5.2.2. Simulasi MARS Birespon Kontinu pada Fungsi Eksponensial-Trigonometri	107
5.2.3. Simulasi MARS Birespon Kontinu pada Fungsi Polinomial-Eksponensial	115
 BAB 6. PEMODELAN PERSENTASE PENDUDUK MISKIN DAN IPM DENGAN MARS BIRESPOK KONTINU.....	 125
6.1. Eksplorasi Data Persentase Penduduk Miskin, IPM dan Variabel yang Mempengaruhinya	125
6.2. Model Persentase Penduduk Miskin dan IPM	132
6.3. Validasi Model Persentase Penduduk Miskin dan IPM	141
 BAB 7. KESIMPULAN DAN KEBERLANJUTAN	 143
7.1. Kesimpulan	143
7.2. Keberlanjutan	147
 DAFTAR PUSTAKA	 149
LAMPIRAN	157
DAFTAR RIWAYAT HIDUP	207

DAFTAR GAMBAR

Gambar 1.1.	Letak Orisinalitas Penelitian	10
Gambar 2.1.	Ilustrasi Pola Hubungan antara 1 Respon dengan 1 Prediktor..	19
Gambar 2.2.	Ilustrasi Pola Nilai GCV pada Berbagai Posisi Titik Knot dalam Estimator <i>Truncated Spline</i>	20
Gambar 2.3.	Ilustrasi Letak Titik Knot Optimal dari Estimasi Fungsi Regresi <i>Truncated Spline</i>	21
Gambar 2.4.	Ilustrasi Pola Hubungan antara 1 Respon dan 2 Prediktor	29
Gambar 2.5.	Ilustrasi Pola Nilai err pada Penambahan Pasangan Fungsi Basis ke-1 (B_1 dan B_2) dalam Estimator MARS	31
Gambar 2.6.	Ilustrasi Pola Nilai GCV pada Penghapusan Fungsi Basis dalam <i>Forward Stepwise</i> dalam Estimator MARS	32
Gambar 2.7.	Ilustrasi Grafik Estimasi Fungsi Regresi MARS	34
Gambar 5.1.	Bentuk Kurva Regresi pada Simulasi Fungsi Asli MARS	80
Gambar 5.2.	<i>Scatter Plot</i> Respon dengan Prediktor pada Simulasi Fungsi Asli MARS	80
Gambar 5.3.	Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Asli MARS dengan Fungsi Basis Optimal	83
Gambar 5.4.	Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Asli MARS dengan Sedikit Fungsi Basis	84
Gambar 5.5.	Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Asli MARS dengan Banyak Fungsi Basis	84
Gambar 5.6.	<i>Boxplot</i> Estimasi Parameter untuk Berbagai Ukuran Sampel pada Simulasi Fungsi Asli MARS pada Respon 1	86
Gambar 5.7.	<i>Boxplot</i> Estimasi Parameter untuk Berbagai Ukuran Sampel pada Simulasi Fungsi Asli MARS pada Respon 2	87
Gambar 5.8.	<i>Boxplot</i> Estimasi Titik Knot untuk Berbagai Ukuran Sampel pada Simulasi Fungsi Asli MARS pada Respon 1	88
Gambar 5.9.	<i>Boxplot</i> Estimasi Titik Knot untuk Berbagai Ukuran Sampel pada Simulasi Fungsi Asli MARS pada Respon 2	88

Gambar 5.10. <i>Boxplot</i> Estimasi Varians <i>Error</i> pada Berbagai Ukuran Sampel pada Simulasi Fungsi Asli MARS	89
Gambar 5.11. <i>Boxplot</i> Estimasi Korelasi <i>Error</i> ada Berbagai Ukuran Sampel pada Simulasi Fungsi Asli MARS	89
Gambar 5.12. <i>Boxplot</i> Estimasi Parameter untuk Berbagai Varians <i>Error</i> pada Simulasi Fungsi Asli MARS pada Respon 1	90
Gambar 5.13. <i>Boxplot</i> Estimasi Parameter untuk Berbagai Varians <i>Error</i> pada Simulasi Fungsi Asli MARS pada Respon 2	91
Gambar 5.14. <i>Boxplot</i> Estimasi Titik Knot untuk Berbagai Varians <i>Error</i> pada Simulasi Fungsi Asli MARS pada Respon 1	92
Gambar 5.15. <i>Boxplot</i> Estimasi Titik Knot untuk Berbagai Varians <i>Error</i> pada Simulasi Fungsi Asli MARS pada Respon 2	92
Gambar 5.16. <i>Boxplot</i> Estimasi Varians <i>Error</i> untuk Berbagai Ukuran Varians <i>Error</i> pada Simulasi Fungsi Asli MARS	93
Gambar 5.17. <i>Boxplot</i> Estimasi Korelasi <i>Error</i> untuk Berbagai Ukuran Varians <i>Error</i> pada Simulasi Fungsi Asli MARS	93
Gambar 5.18. <i>Boxplot</i> Rasio <i>RMSEP</i> Birespon berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Asli MARS	97
Gambar 5.19. <i>Boxplot</i> Rasio <i>RMSEP</i> berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Asli MARS pada Setiap Respon	97
Gambar 5.20. Rata-rata Rasio <i>RMSEP</i> Birespon berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Asli MARS	98
Gambar 5.21. Rata-rata Rasio <i>RMSEP</i> berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Asli MARS pada Setiap Respon	98
Gambar 5.22. Bentuk Kurva Regresi pada Simulasi Fungsi Trigonometri-Polinomial	100
Gambar 5.23. <i>Scatter Plot</i> Respon dengan Prediktor pada Simulasi Fungsi Trigonometri-Polinomial	101
Gambar 5.24. Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Trigonometri-Polinomial dengan Fungsi Basis Optimal	102
Gambar 5.25. Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Trigonometri-Polinomial dengan Sedikit Fungsi Basis	103

Gambar 5.26. Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Trigonometri-Polinomial dengan Banyak Fungsi	104
Gambar 5.27. <i>Boxplot</i> Rasio <i>RMSEP</i> Birespon berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Trigonometri-Polinomial	105
Gambar 5.28. <i>Boxplot</i> Rasio <i>RMSEP</i> berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Trigonometri-Polinomial pada Setiap Respon	105
Gambar 5.29. Rata-rata Rasio <i>RMSEP</i> Birespon berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Trigonometri-Polinomial	106
Gambar 5.30. Rata-rata Rasio <i>RMSEP</i> berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Trigonometri-Polinomial pada Setiap Respon	106
Gambar 5.31. Bentuk Kurva Regresi pada Simulasi Fungsi Eksponensial-Trigonometri	107
Gambar 5.32. <i>Scatter Plot</i> Respon dengan Prediktor pada Simulasi Fungsi Eksponensial-Trigonometri	108
Gambar 5.33. Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Eksponensial-Trigonometri dengan Fungsi Basis Optimal	110
Gambar 5.34. Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Eksponensial-Trigonometri dengan Sedikit Fungsi	111
Gambar 5.35. Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Eksponensial-Trigonometri dengan Banyak Fungsi Basis.....	112
Gambar 5.36. <i>Boxplot</i> Rasio <i>RMSEP</i> Birespon berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Eksponensial-Trigonometri	112
Gambar 5.37. <i>Boxplot</i> Rasio <i>RMSEP</i> berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Eksponensial-Trigonometri pada Setiap Respon	113
Gambar 5.38. Rata-rata Rasio <i>RMSEP</i> Birespon berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Eksponensial-Trigonometri	114

Gambar 5.39. Rata-rata Rasio <i>RMSEP</i> berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Eksponensial-Trigonometri pada Setiap Respon	114
Gambar 5.40. Bentuk Kurva Regresi pada Simulasi Fungsi Polinomial-Eksponensial	115
Gambar 5.41. <i>Scatter Plot</i> Respon dengan Prediktor pada Simulasi Fungsi Polinomial-Eksponensial	116
Gambar 5.42. Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Polinomial-Eksponensial dengan Fungsi Basis Optimal	118
Gambar 5.43. Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Polinomial-Eksponensial dengan Sedikit Fungsi Basis	119
Gambar 5.44. Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Polinomial-Eksponensial dengan Banyak Fungsi	119
Gambar 5.45. <i>Boxplot</i> Rasio <i>RMSEP</i> Birespon berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Polinomial-Eksponensial	120
Gambar 5.46. <i>Boxplot</i> Rasio <i>RMSEP</i> berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Polinomial-Eksponensial pada Setiap Respon	121
Gambar 5.47. Rata-rata Rasio <i>RMSEP</i> Birespon berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Polinomial-Eksponensial	121
Gambar 5.48. Rata-rata Rasio <i>RMSEP</i> berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Polinomial-Eksponensial pada Setiap Respon	122
Gambar 6.1. <i>Scatter Plot</i> Respon dengan Prediktor pada Data Pemodelan Persentase Penduduk Miskin dan IPM di Pulau Jawa	128
Gambar 6.2. Plot Uji Q-Q Plot pada Residual Model	134
Gambar 6.3. Bentuk Fungsi Basis BF_{11} dan BF_{12}	135
Gambar 6.4. Bentuk Fungsi Basis BF_{12}	135
Gambar 6.5. Bentuk Fungsi Basis BF_{11}	139

DAFTAR TABEL

Tabel 2.1. Perbandingan Nilai GCV berdasarkan Banyaknya Titik Knot dalam Estimator <i>Truncated Spline</i>	20
Tabel 2.2. Penambahan Fungsi Basis pada <i>Forward Stepwise</i> dalam Estimator MARS	30
Tabel 2.3. Penghapusan Fungsi Basis pada <i>Backward Stepwise</i> dalam Estimator MARS	32
Tabel 3.1. Struktur Data Penelitian	57
Tabel 5.1. Estimasi Fungsi Basis Optimal pada Simulasi Fungsi Asli MARS pada Respon 1	81
Tabel 5.2. Estimasi Fungsi Basis Optimal pada Simulasi Fungsi Asli MARS pada Respon 2	81
Tabel 5.3. Estimasi Parameter Model MARS Birespon Kontinu pada Simulasi Fungsi Asli MARS	82
Tabel 5.4. Jumlah Estimasi Model yang Sesuai dengan Model Aslinya pada Simulasi Fungsi Asli MARS dengan Replikasi 100 kali	95
Tabel 5.5. Estimasi Parameter dan Fungsi Basis Optimal MARS Birespon pada Simulasi Fungsi Trigonometri-Polinomial.....	102
Tabel 5.6. Estimasi Parameter dan Fungsi Basis Optimal MARS Birespon pada Simulasi Fungsi Eksponensial-Trigonometri.....	109
Tabel 5.7. Estimasi Parameter dan Fungsi Basis Optimal MARS Birespon pada Simulasi Fungsi Polinomial-Eksponensial	117
Tabel 6.1. Statistik Deskriptif Variabel Penelitian pada Data Riil	126
Tabel 6.2. Uji Non-Linieritas <i>Ramsey Reset Test</i> pada Data Riil	132
Tabel 6.3. Estimasi Parameter dan Fungsi Basis Optimal pada Pemodelan Persentase Penduduk Miskin dan IPM di Pulau Jawa	133
Tabel 6.4. Perbandingan Ketepatan Estimator Model MARS Birespon Kontinu dengan Model Parametrik pada Pemodelan Data Riil	140
Tabel 6.5. Perbandingan Nilai <i>RMSECV</i> dari Model yang Melibatkan Korelasi dengan Tidak Melibatkan Korelasi pada Data Riil	142

DAFTAR LAMPIRAN

Lampiran 1. Syntax untuk Membangkitkan Data Simulasi pada Percobaan Fungsi Asli MARS	157
Lampiran 2. Syntax untuk Membangkitkan Data Simulasi pada Kombinasi Fungsi Percobaan Trigonometri-Polinomial	159
Lampiran 3. Syntax untuk Membangkitkan Data Simulasi pada Kombinasi Fungsi Percobaan Eksponensial-Trigonometri	161
Lampiran 4. Syntax untuk Membangkitkan Data Simulasi pada Kombinasi Fungsi Percobaan Polinomial-Eksponensial	163
Lampiran 5. Syntax Pemodelan MARS Birespon Kontinu pada Data Simulasi	165
Lampiran 6. Syntax untuk Menghitung Rasio $RMSEP$ pada Data Simulasi.....	169
Lampiran 7. Syntax Pemodelan Persentase Penduduk Miskin dan IPM Menggunakan Estimator MARS Birespon Kontinu	174
Lampiran 8. Syntax untuk Menghitung Nilai $RMSECV$ pada Pemodelan Persentase Penduduk Miskin dan IPM Menggunakan Estimator MARS Birespon Kontinu	177
Lampiran 9. Syntax untuk Uji Linieritas Data	180
Lampiran 10. Syntax untuk Uji Normal Multivariat Data	182
Lampiran 11. Data Simulasi pada Fungsi Asli MARS	184
Lampiran 12. Data Simulasi pada Fungsi Percobaan Trigonometri-Polinomial	188
Lampiran 13. Data Simulasi pada Fungsi Percobaan Eksponensial-Trigonometri	192
Lampiran 14. Data Simulasi pada Fungsi Percobaan Polinomial-Eksponensial	196
Lampiran 15. Data Persentase Penduduk Miskin (y_1), Indeks Pembangunan Manusia (y_2), Pertumbuhan Ekonomi (x_1), Angka Beban Ketergantungan (x_2), Persentase Pengeluaran Perkapita untuk Makanan (x_3), Tingkat Pengangguran Terbuka (x_4) dan Persentase Penduduk yang Pekerjaan Utamanya di Sektor Pertanian (x_5) menurut Kabupaten/Kota di Pulau Jawa	200
Lampiran 16. GUI MARS Birespon Kontinu	204

DAFTAR LAMBANG DAN ARTI

$y_{\ell i}$: respon ke- ℓ dan observasi ke- i , untuk $\ell=1,2$ dan $i=1,2,\dots,n$.

x_{ji} : prediktor ke- j dan observasi ke- i , untuk $j=1,2,\dots,p$ dan $i=1,2,\dots,n$.

$\varepsilon_{\ell i}$: *random error* pada respon ke- ℓ dan observasi ke- i , untuk $\ell=1,2$ dan $i=1,2,\dots,n$.

n : banyaknya observasi / pengamatan.

p : banyaknya prediktor.

f_{ℓ} : fungsi regresi nonparametrik yang menunjukkan pola hubungan antara prediktor terhadap respon ke- ℓ .

$E(\varepsilon)$: rata-rata / nilai harapan dari ε .

$\text{Var}(\varepsilon)$: varians ε .

α : vektor parameter/koeffisien regresi MARS Birespon Kontinu.

$\hat{\alpha}$: estimasi dari α .

σ_{11} : varians *random error* pada respon 1.

σ_{22} : varians *random error* pada respon 2.

σ_{12} : kovarians *random error* antar respon 1 dan respon 2.

ρ : korelasi *random error* antar respon 1 dan respon 2.

$\mathbf{B}(\mathbf{x}, t)$: matriks fungsi basis dalam estimator MARS Birespon Kontinu.

$t_{k_{\ell} m_{\ell}}$: titik knot pada interaksi ke- k_{ℓ} dan fungsi basis ke- m_{ℓ} dalam estimator MARS Birespon Kontinu, dengan $\ell=1,2$ adalah indeks yang menunjukkan respon ke- ℓ .

K : banyaknya interaksi dalam estimator MARS.

M : banyaknya fungsi basis dalam estimator MARS.

\mathbf{W} : matriks bobot dalam estimator MARS Birespon Kontinu.

\mathbf{H} : matriks *hat*.

\mathbf{I} : matriks identitas.

$\sim N$: distribusi *Normal*.

$\sim U$: distribusi *Uniform*.

$L(\cdot)$: fungsi *likelihood*.

$\ln L(\cdot)$: ln dari fungsi *likelihood* (fungsi *log likelihood*).

\otimes : *kroncker product*.

\sum_m : penjumlahan yang meliputi indeks m .

\prod_k : perkalian yang meliputi indeks k .

$\|\cdot\|$: norm / panjang vektor.

\mathbf{A}' : transpose dari matriks \mathbf{A} .

\mathbf{A}^{-1} : invers dari matriks \mathbf{A} .

$tr(\mathbf{A})$: trace dari matriks \mathbf{A} .

$|\mathbf{A}|$: determinan dari matriks \mathbf{A} .

■ : Akhir Bukti.

RMSE : *Root Mean Square Error*.

RMSEP : *Root Mean Square Error Prediction*.

RMSECV : *Root Mean Square Error of Cross Validation*.

GCV : *Generalized Cross Validation*.

R^2 : Koefisien Determinasi.

MLE : *Maximum Likelihood Estimation*.

OLS : *Ordinary Least Square*.

Halaman ini sengaja dikosongkan

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Analisis regresi merupakan suatu metode yang biasa digunakan untuk memodelkan sebuah persoalan riil ke dalam bentuk persamaan matematis yang dapat menjelaskan pola hubungan antara respon dengan prediktor. Pola hubungan tersebut digambarkan dalam kurva regresi, yang dapat diduga dengan tiga jenis pendekatan, yaitu pendekatan regresi parametrik, regresi nonparametrik dan regresi semiparametrik (gabungan dari regresi parametrik dan regresi nonparametrik). Pendekatan regresi parametrik merupakan metode yang paling sering digunakan untuk menyelesaikan pemodelan regresi, dengan syarat bentuk kurva regresi diasumsikan diketahui bentuknya. Pendekatan regresi parametrik umumnya terdiri dari dua kelompok model, yaitu model regresi linier dan model regresi nonlinier.

Pemodelan regresi pada kasus tertentu seringkali tidak dapat diselesaikan dengan pendekatan regresi parametrik dan harus diselesaikan dengan pendekatan regresi nonparametrik yaitu pada kasus-kasus dimana bentuk kurva regresi tidak/belum diketahui bentuknya (Hardle, 1990). Kasus tersebut dapat diselesaikan karena regresi nonparametrik memiliki kemampuan dalam mencari bentuk pola kurva regresi yang tidak/belum diketahui bentuknya. Kemampuan ini didukung oleh adanya parameter pada setiap jenis metode regresi nonparametrik yang membuat estimasi pola kurva regresi menjadi lebih fleksibel. Parameter tersebut diantaranya adalah knot dalam metode MARS (Friedman, 1991; Lewis dan Stevens, 1990; Otok, 2008) dan *truncated* spline (Smith, 1979; Crainiceanu, Ruppert dan Wand, 2004; Budiantara, 2006; Sudiarsa, Budiantara, Suhartono dan Purnami, 2015; Ruliana, Budiantara, Otok dan Wibowo, 2015), parameter penghalus dalam *smoothing* spline (Eubank, 1999; Wang, Guo dan Brown, 2000; Howell, 2007) dan deret fourier (Pane, Budiantara, Zain dan Otok, 2014) serta *bandwith* dalam polinomial lokal (Astuti, Budiantara, Sunaryo dan Dokhi, 2013) dan kernel (Kayri dan Zirhlioglu, 2009).

Penelitian regresi nonparametrik telah banyak dikembangkan baik berdasarkan jenis respon maupun berdasarkan banyaknya respon yang terlibat dalam pemodelan regresi. Penelitian regresi berdasarkan jenis respon dibedakan kedalam model regresi respon kontinu (kuantitatif) dan kategori (kualitatif). Respon kontinu adalah jika skala pengukurannya interval atau rasio, sedangkan respon kategori adalah jika skala pengukurannya nominal atau ordinal. Penelitian regresi berdasarkan banyaknya respon dibedakan menjadi model regresi respon tunggal dan multirespon.

Model regresi respon tunggal merupakan model regresi yang paling umum, yaitu terdiri atas satu respon yang dipengaruhi oleh satu atau beberapa prediktor. Para peneliti yang menerapkan model ini pada regresi nonparametrik respon kontinu diantaranya adalah Wang (1998) yang menerapkan *smoothing spline* pada *error* berkorelasi. Kim dan Gu (2004) yang menerapkan *smoothing spline* untuk respon Gaussian. Aydin (2007) yang menunjukkan kebaikan *smoothing spline* dibanding *smoothing kernel* untuk pemodelan *Gross National Product* (GNP) di Turki. Nottingham dan Cook (2001) yang menerapkan lokal linier untuk estimasi data *time series*. Hallin, Lu dan Tran (2004) yang menerapkan lokal linier pada regresi spasial. Liang dan Chen (2005) yang membandingkan polinomial lokal dengan lokal linier untuk mengestimasi tekanan *Esophageal* pada penyakit *Gastroesophageal Reflux*. Su dan Ullah (2008) yang menerapkan polinomial lokal pada model persamaan simultan. Penelitian regresi nonparametrik respon tunggal kategori telah banyak dikembangkan pula, diantaranya adalah Ripley (1994) yang mengkaji metode *Neural Network* (NN) untuk klasifikasi, yang diterapkan pada *synthetic data*, *sonar data* dan *forensic glass data*. Xiong dan Meullenet (2002) yang menggunakan metode *Multivariate Adaptive Regression Spline* (MARS) pada pemodelan efek aktifitas air, pH dan Potassium pada pertumbuhan *Sacharomyces Cerevisiac*. Peneliti lainnya yaitu Otok (2005) yang menerapkan MARS untuk klasifikasi bank di Indonesia.

Model regresi multirespon berbeda halnya dengan model regresi respon tunggal, dimana model ini terdiri dari beberapa persamaan dengan asumsi terdapat korelasi antar respon. Para peneliti umumnya mengakomodir adanya korelasi ini dengan menggunakan matriks kovarians yang digunakan sebagai matriks

pembobot pada saat estimasi parameter model. Ada beberapa peneliti yang telah mengkaji pemodelan regresi birespon dan regresi multirespon pada beberapa metode regresi nonparametrik untuk respon kontinu diantaranya adalah Wang *et al.* (2000) yang mengkaji estimator spline untuk data birespon yang diterapkan pada data hormon ACTH dan hormon Cortisol. Welsh dan Yee (2006) yang mengkaji regresi nonparametrik birespon pada estimator lokal linier dan menerapkannya pada data tekanan darah. Lestari, Budiantara, Sunaryo dan Mashuri (2012) yang mengkaji estimasi regresi nonparametrik multirespon berdasarkan *smoothing* spline untuk kasus heteroskedastisitas. Chamidah, Budiantara, Sunaryo dan Zain (2012) yang mengkaji regresi multirespon dengan pendekatan polinomial lokal dan menerapkannya pada data pertumbuhan balita. Wibowo, Haryatmi dan Budiantara (2013) yang mengkaji model regresi semiparametrik multirespon pada estimator *smoothing* spline dan menerapkannya pada kasus kinerja aktifitas perbankan di setiap provinsi di Indonesia. Penelitian terakhir dilakukan oleh Fernandes, Budiantara, Otok dan Suhartono (2014) yang mengkaji regresi nonparametrik birespon pada estimator *smoothing* spline untuk data longitudinal dan penerapannya pada data penderita TB Paru di Kota Malang.

Penelitian-penelitian model regresi birespon/multirespon kontinu pada beberapa jenis metode regresi nonparametrik yang disebutkan di atas (*smoothing* spline, lokal linier dan polinomial lokal) umumnya masih menggunakan model aditif antar prediktor dalam pemodelan datanya. Peneliti-peneliti terdahulu belum ada yang mengembangkan model regresi nonparametrik multirespon yang melibatkan interaksi antar prediktor dalam pemodelannya. Salah satu metode regresi nonparametrik yang mampu mengakomodir pengaruh aditif maupun pengaruh interaksi antara prediktor dalam pemodelan datanya adalah MARS (Friedman, 1991). Kemampuan ini memungkinkan MARS untuk mendapatkan hasil estimasi yang lebih baik atau mendekati bentuk kurva regresi dari pola hubungan respon dan prediktor yang tidak diketahui bentuknya, dibandingkan dengan hanya menggunakan model aditif, diukur dengan kriteria *generalized cross validation* (GCV). Hastie, Tibshirani, Friedman (2008) menyatakan bahwa MARS adalah metode regresi adaptif yang umumnya baik digunakan pada pemodelan regresi nonparametrik yang melibatkan banyak prediktor, meskipun

demikian menurut Friedman (1991) metode MARS juga memiliki kemampuan yang sama baiknya untuk pemodelan yang melibatkan sedikit prediktor (satu atau dua prediktor). Hal ini ditunjukkan Friedman (1991) dalam simulasi yang melibatkan dua prediktor. Hasil simulasi menunjukkan bahwa grafik estimasi kurva regresi dan nilai *expected square error* dari MARS sama baiknya dengan hasil estimasi yang diperoleh dari metode regresi nonparametrik nonadaptif, dimana dalam simulasi itu dibandingkan dengan metode spline.

Metode MARS pertama kali diperkenalkan oleh Friedman pada tahun 1988. Metode ini merupakan pendekatan regresi nonparametrik yang merupakan kombinasi antara metode *truncated spline* dengan *recursive partitioning regression* (RPR). Metode MARS mampu menangani data yang perilakunya berubah-ubah pada sub-sub interval tertentu, hal ini terjadi karena dalam MARS terdapat titik knot, yaitu titik yang menunjukkan terjadinya perubahan pola perilaku data. Metode MARS dibangun dengan algoritma *stepwise* yang terdiri dari *forward stepwise* dan *backward stepwise*. *Forward stepwise* membangun model dengan menambahkan fungsi basis *truncated spline* (knot dan interaksi) hingga diperoleh model dengan jumlah fungsi basis maksimum dan *backward stepwise* digunakan untuk mendapatkan model parsimoni dengan memilih fungsi basis *forward stepwise* yang kontribusinya paling signifikan terhadap dugaan respon berdasarkan nilai GCV minimum.

Pemodelan MARS telah banyak dilakukan baik pada regresi dengan respon kontinu maupun kategorik, dan umumnya dikaji pada pemodelan regresi respon tunggal. Pemodelan tersebut diantaranya dilakukan oleh Bakin, Hegland dan Osborne (1997) yang mengkaji pengembangan metode MARS menggunakan basis B-Spline pada data simulasi. Denison, Mallick dan Smith (1998) yang mengkaji pemodelan Bayesian MARS dan diterapkan pada data Ozon. Sephton (2001) yang mengkaji kemampuan variabel finansial untuk peramalan resesi di Amerika menggunakan metode MARS. Abraham dan Steinberg (2001) yang membandingkan metode MARS dan NN pada pemodelan data *gas furnace time series*. Xu, Massart, Liang dan Fang (2003) yang mengkaji peramalan penjualan produk teknologi informasi menggunakan model *hybrid* MARS dan *Support Vector Regression* (SVR). Otok, Guritno dan Subanar (2006) yang menerapkan

metode MARS pada pemodelan volume perdagangan saham. Penelitian lainnya adalah Mina (2008) yang menerapkan metode MARS untuk pengklasifikasian penduduk miskin di Philipina.

Pemodelan MARS yang melibatkan lebih dari satu respon pernah dikembangkan pula oleh beberapa peneliti diantaranya adalah Milborrow (2015) dan Jekabsons (2015) dalam bentuk *multivariate response*, dan membuat aplikasinya masing-masing dalam *R-package (earth)* dan matlab (ARESLab). Hanya saja pengembangan model MARS yang dilakukan oleh kedua peneliti tersebut belum menyelesaikan permasalahan regresi nonparametrik multirespon yang ditemui pada peneliti sebelumnya (Wang *et al.*, 2000, Welsh *et al.*, 2006, Lestari *et al.*, 2012, Chamidah *et al.*, 2012, Wibowo *et al.*, 2013 dan Fernandez *et al.*, 2014), dimana permasalahan peneliti regresi nonparametrik multirespon sebelumnya tidak mengakomodir pengaruh interaksi dalam pemodelannya. Hal ini disebabkan karena rancangan asumsi model regresi yang dikembangkan oleh Milborrow (2015) dan Jekabsons (2015) berbeda dengan rancangan asumsi model regresi nonparametrik multirespon yang dikembangkan oleh Wang *et al.* (2000), Welsh *et al.* (2006), Lestari *et al.* (2012), Chamidah *et al.* (2012), Wibowo *et al.* (2013) dan Fernandez *et al.* (2014).

Rancangan model yang dikembangkan oleh Milborrow (2015) dan Jekabsons (2015) pada dasarnya mengikuti bentuk *multivariate regression* yang mengoptimisasi pemodelan semua respon secara simultan menggunakan MARS, sehingga setiap model respon yang diperoleh memiliki fungsi basis yang sama tetapi dengan koefisien regresi yang berbeda. Menurut Milborrow (2015) dan Jekabsons (2015), model tersebut baik untuk tujuan interpretasi karena perubahan satu fungsi basis MARS diikuti oleh perubahan lebih dari satu respon secara bersama-sama, namun dampaknya kurang akurat untuk estimasi karena fungsi basis yang dihasilkan tidak fleksibel untuk setiap respon. Hal ini dinyatakan oleh kriteria nilai *generalization R-square* (GRSq) dari model tersebut yang tidak sebaik jika dimodelkan secara independen. Model tersebut juga terbatas digunakan pada prediktor yang sama untuk semua respon.

Penelitian disertasi ini berfokus pada pengembangan model MARS birespon kontinu yang mengakomodir adanya asumsi korelasi antar respon.

Asumsi ini diakomodir menggunakan matriks pembobot pada saat estimasi parameter model. Pada dasarnya pengembangan model yang dilakukan dalam disertasi ini memiliki perbedaan dengan penelitian Milborrow (2015) dan Jekabsons (2015), yang disebabkan oleh perbedaan rancangan asumsi model yang dikembangkan. Model yang dihasilkan pun akan berbeda, karena fungsi basis yang dihasilkan dalam disertasi ini fleksibel untuk setiap respon. Selain itu, penggunaannya juga tidak dibatasi hanya pada prediktor yang sama untuk semua respon.

Pada penelitian ini, model regresi nonparametrik birespon kontinu dengan pendekatan MARS akan diterapkan pada pemodelan data kesejahteraan rakyat yang melibatkan respon persentase penduduk miskin dan indeks pembangunan manusia (IPM). Beberapa peneliti telah mengkaji pemodelan dari masing-masing respon tersebut seperti Wulandari dan Budiantara (2014) yang mengkaji pemodelan persentase penduduk miskin di Jawa Timur yang dipengaruhi pengeluaran perkapita makanan dengan pendekatan spline. Pratowo (2011) mengkaji pemodelan IPM yang dipengaruhi pola konsumsi non makanan dan angka beban ketergantungan. Son dan Kakwani (2004) serta Ijaiya, Bello dan Ajayi (2011) mengkaji pemodelan kemiskinan beberapa negara di Asia dan Nigeria yang dipengaruhi pertumbuhan ekonomi. Ranis, Stewart dan Ramirez (2000) mengkaji pemodelan IPM yang dipengaruhi pertumbuhan ekonomi pada beberapa negara berkembang di dunia. Budiantara, Diana, Purhadi dan Darmesto (2011) mengkaji pemodelan kemiskinan di Indonesia yang dipengaruhi angka pengangguran dengan pendekatan bayesian spline. Nurbaeti (2013) yang mengkaji pemodelan IPM yang dipengaruhi angka pengangguran di provinsi Jawa Tengah. Penelitian lainnya yaitu Ayomi (2014) yang mengkaji pengaruh penyerapan sektor tenaga kerja terhadap tingkat kemiskinan dan IPM.

Penelitian-penelitian yang disebutkan di atas umumnya memodelkan data tingkat kesejahteraan rakyat dengan masih melibatkan satu respon, baik untuk pemodelan respon persentase penduduk miskin maupun untuk pemodelan respon IPM. Beberapa kajian literatur (Hidayat, 2008; Madan, 2012; BPS, 2013; Tarigan, 2014; Pudjianto dan Syawie, 2015) menyatakan bahwa antara kedua respon tersebut saling berkorelasi. Kajian tersebut menyatakan bahwa terdapat korelasi

negatif antara kemiskinan dengan IPM, dimana semakin rendah persentase penduduk miskin di suatu wilayah maka semakin tinggi nilai IPM di wilayah tersebut. Hal ini berarti bahwa pemodelan data tingkat kesejahteraan rakyat (persentase penduduk miskin dan IPM) seharusnya dapat dilakukan dengan pemodelan birespon yang melibatkan korelasi dari kedua respon tersebut.

Penelitian awal dalam disertasi ini dimulai dengan kajian teori untuk mendapatkan estimasi fungsi regresi dalam model MARS birespon kontinu. Kajian teori juga dilakukan untuk mendapatkan matriks bobot untuk mengakomodir adanya korelasi antar respon dalam estimator MARS birespon kontinu. Kajian berikutnya adalah menerapkan hasil kajian teoritik pada data simulasi untuk mengetahui kebaikan model yang diperoleh. Kajian akhir dilakukan dengan menerapkan model MARS birespon kontinu pada data riil yaitu pada pemodelan data tingkat kesejahteraan rakyat di pulau Jawa.

1.2. Permasalahan

Penelitian dalam analisis regresi seringkali tidak hanya ditemui pada pemodelan hubungan antara satu respon dengan satu atau beberapa prediktor, tetapi dapat ditemui pula pada pemodelan regresi birespon yang saling berkorelasi. Hal ini berarti dalam pemodelan regresi birespon tidak hanya memodelkan bagaimana hubungan prediktor terhadap respon, tetapi juga mengakomodir hubungan antar respon. Penelitian MARS yang dikembangkan para peneliti selama ini lebih banyak untuk pemodelan regresi satu respon baik untuk respon kontinu maupun kategori, seperti dilakukan oleh Bakin *et al.* (1997), Denisson *et al.* (1998), Sephton (2001), Abraham *et al.* (2001), Xu *et al.* (2003), Otok *et al.* (2006) dan Mina *et al.* (2008). Penelitian MARS birespon kontinu yang mengakomodir adanya korelasi antar respon belum dikembangkan oleh para peneliti, sehingga dapat dilakukan dalam penelitian disertasi ini. Permasalahan statistik yang muncul dalam penelitian ini adalah bagaimana mendapatkan estimator MARS birespon kontinu dengan mengakomodir adanya korelasi diantara kedua respon tersebut.

Adanya korelasi antar respon dalam estimator MARS birespon kontinu diakomodir melalui matriks bobot pada saat estimasi parameter model. Matriks

bobot secara teori diasumsikan sebagai *fixed variable*, padahal dalam kasus riil tidak diketahui nilainya. Hal ini berarti bahwa matriks bobot tersebut harus diestimasi dari data. Permasalahan statistik yang muncul dalam penelitian ini adalah bagaimana mengestimasi matriks bobot dengan menggunakan fungsi basis MARS.

Penerapan model MARS birespon dalam penelitian ini dilakukan pada data simulasi dan data riil. Penerapan pada data riil dilakukan untuk memodelkan data tingkat kesejahteraan rakyat yang melibatkan dua respon yaitu persentase penduduk miskin dan IPM. Beberapa literatur (Hidayat, 2008; Madan, 2012; BPS, 2013; Tarigan, 2014; Pudjianto dan Syawie, 2015) menyatakan bahwa terdapat hubungan diantara kedua respon tersebut. Permasalahan statistik dari sisi simulasi adalah bagaimana merancang studi simulasi dari model MARS birespon kontinu untuk melihat kebaikan dari model yang diperoleh. Permasalahan statistik dari sisi data riil adalah bagaimana memodelkan tingkat kesejahteraan rakyat dengan mengakomodir korelasi antara respon persentase penduduk miskin dan IPM.

1.3. Tujuan Penelitian

Tujuan penelitian yang ingin dicapai untuk menyelesaikan permasalahan di atas adalah sebagai berikut :

1. Mendapatkan estimator regresi nonparametrik birespon kontinu menggunakan fungsi MARS.
2. Mendapatkan estimasi bobot yang mengakomodir adanya korelasi antar respon dalam estimator MARS birespon kontinu.
3. Merancang studi simulasi pada model MARS birespon kontinu untuk melihat kebaikan dari model yang diperoleh.
4. Menerapkan model MARS birespon kontinu pada pemodelan data tingkat kesejahteraan rakyat yang melibatkan dua respon yaitu persentase penduduk miskin dan IPM.

1.4. Manfaat Penelitian

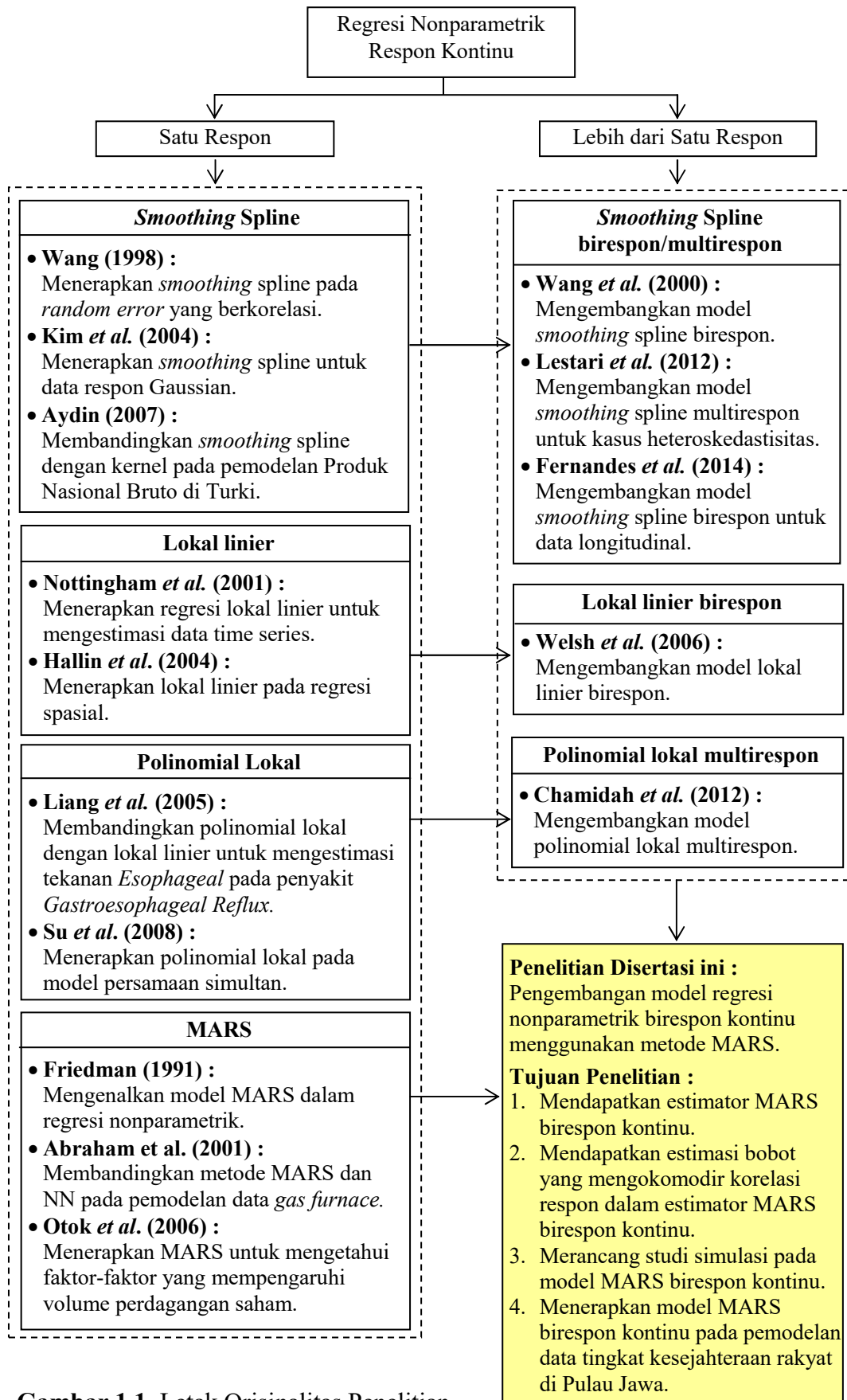
Manfaat yang dapat diambil dari penelitian ini terbagi atas dua sisi yaitu dari sisi statistik dan sisi terapan. Manfaat penelitian dari sisi statistik adalah

memberikan kontribusi teori bagi pengembangan model regresi nonparametrik birespon kontinu menggunakan estimator MARS. Manfaat penelitian dari sisi terapan adalah memberikan kontribusi model baru sebagai alternatif dalam mengestimasi tingkat kesejahteraan rakyat di Pulau Jawa, yang diestimasi dari angka persentase penduduk miskin dan angka IPM. Sehingga diharapkan setiap kebijakan yang akan dilakukan pemerintah untuk kesejahteraan rakyat dapat dievaluasi dari hasil estimasi model ini.

1.5. Orisinalitas Penelitian

Penelitian regresi multirespon yang mengakomodir adanya korelasi antar respon telah banyak dikaji oleh para peneliti pada berbagai jenis fungsi regresi nonparametrik, seperti pada spline (Wang *et al.*, 2000; Lestari *et al.*, 2012; Wibowo *et al.*, 2013; Fernandez *et al.*, 2014), lokal linier (Welsh *et al.*, 2006) dan polinomial lokal (Chamidah *et al.*, 2012). Penelitian regresi nonparametrik multirespon yang dilakukan oleh para peneliti tersebut, umumnya masih menggunakan model aditif, sedangkan yang menggunakan model interaksi antar prediktor belum dilakukan oleh para peneliti hingga saat ini. Hal ini yang menjadi ide orisinalitas untuk mengembangkan model MARS birespon kontinu dalam disertasi ini, mengingat metode MARS memiliki kemampuan untuk mengakomodir pengaruh aditif maupun pengaruh interaksi antar prediktor dalam pemodelannya.

Orisinalitas berikutnya dalam penelitian ini adalah estimasi matriks bobot yang bertujuan untuk mengakomodir adanya korelasi antar respon dalam pemodelan MARS birespon kontinu. Matriks bobot ini secara teori diasumsikan sebagai *fixed variable*, namun dalam penerapannya tidak diketahui nilainya sehingga harus diestimasi dari data. Orisinalitas lain dalam penelitian ini adalah penerapan model MARS birespon kontinu pada pemodelan kasus data riil, yaitu pada pemodelan data tingkat kesejahteraan rakyat yang melibatkan adanya korelasi antar respon persentase penduduk miskin dan IPM di Pulau Jawa. Hal ini belum pernah dilakukan oleh para peneliti-peneliti lain sampai dengan saat ini. Letak orisinalitas penelitian dalam disertasi ini lebih jelasnya dapat dilihat pada Gambar 1.1.



Gambar 1.1. Letak Orisinalitas Penelitian

1.6. Batasan Penelitian

Batasan penelitian dalam hal ini diperlukan agar penelitian yang dilakukan lebih fokus sesuai dengan rentang waktu yang direncanakan. Batasan-batasan yang akan dikerjakan adalah sebagai berikut:

1. Respon dan prediktor yang digunakan dalam penelitian ini menggunakan skala pengukuran interval atau rasio
2. Pola hubungan antara respon dengan prediktor memiliki bentuk hubungan yang belum diketahui bentuknya.
3. Estimasi parameter dan bobot diperoleh dengan metode *maximum likelihood estimation* (MLE).
4. Metode MARS yang digunakan dalam penelitian ini adalah MARS berderajat satu (linier).
5. Matriks bobot dalam estimator MARS birespon kontinu diasumsikan diketahui (*fixed variable*).
6. Fungsi basis MARS birespon ditetapkan sama dengan fungsi basis MARS respon tunggal yang diperoleh dengan menggunakan kriteria GCV.
7. Maksimum interaksi yang digunakan adalah tiga interaksi.
8. Kriteria kebaikan model yang digunakan adalah *root mean square error* (RMSE).

Halaman ini sengaja dikosongkan

BAB 2

KAJIAN PUSTAKA DAN DASAR TEORI

Bab ini menjelaskan mengenai beberapa konsep, hasil penelitian dan teorema yang menjadi kajian pustaka dan dasar teori untuk menunjang penelitian. Kajian diawali dengan penjelasan teori analisis regresi yang meliputi regresi parametrik dan regresi nonparametrik. Kajian regresi difokuskan pada pembahasan salah satu metode regresi nonparametrik yang digunakan dalam penelitian ini yaitu MARS. Metode MARS merupakan pengembangan dari metode *truncated* spline yang diuraikan pula dalam bab ini. Kajian berikutnya menjelaskan tentang asumsi model regresi nonparametrik birespon/multirespon kontinu, ukuran kesesuaian model dan uji linieritas yang menunjang hasil penelitian ini. Kajian pustaka juga dilengkapi dengan menyebutkan beberapa teorema dan preposisi matriks aljabar yang akan digunakan untuk mendukung penyelesaian tujuan teori penelitian ini. Model MARS birespon kontinu yang dikembangkan dalam penelitian ini selanjutnya akan diterapkan pada pemodelan data tingkat kesejahteraan rakyat yang melibatkan dua respon yaitu persentase penduduk miskin dan IPM serta beberapa prediktor yang mempengaruhinya. Definisi operasional dari masing-masing respon dan prediktor tersebut selanjutnya akan diuraikan pada kajian akhir dari bab ini.

2.1. Analisis Regresi

Analisis regresi dalam statistika digunakan untuk menjelaskan dugaan tentang adanya hubungan kausal antara respon dengan prediktor. Terdapat dua pendekatan yang sering digunakan dalam analisis regresi yaitu regresi parametrik dan nonparametrik. Perbedaan utama dari kedua pendekatan regresi ini terletak pada asumsi bentuk pola kurva regresi yang digunakan dalam pemodelan data.

2.1.1. Model Regresi Parametrik

Regresi parametrik merupakan metode regresi yang digunakan ketika pola hubungan antara respon dan prediktor diketahui bentuknya, seperti linier, kuadratik atau kubik. Salah satu model regresi parametrik yang sering digunakan

adalah model regresi parametrik linier berganda dengan persamaan sebagai berikut (Montgomery, Peck dan Vining, 2012):

$$y_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ji} + \varepsilon_i; \quad i=1,2,\dots,n, \quad (2.1)$$

dengan y_i adalah respon pada pengamatan ke- i , x_{ji} adalah prediktor ke- j pada pengamatan ke- i , β_0 adalah intersep, β_j adalah koefisien regresi ke- j , ε_i adalah *random error* pada pengamatan ke- i , p adalah banyaknya prediktor dan n adalah banyaknya pengamatan. *Random error* (ε_i) dalam persamaan (2.1) diasumsikan independen berdistribusi normal dengan *mean* nol dan varians σ^2 .

Persamaan (2.1) dengan n data pengamatan dapat dinyatakan dalam bentuk matriks sebagai berikut:

$$\underline{y} = \mathbf{X}\underline{\beta} + \underline{\varepsilon}, \quad (2.2)$$

dengan \underline{y} adalah vektor respon berorde $n \times 1$, \mathbf{X} adalah matriks prediktor berorde $n \times (p+1)$, $\underline{\beta}$ adalah vektor dari koefisien regresi berorde $(p+1) \times 1$, dan $\underline{\varepsilon}$ adalah vektor *random error* berorde $n \times 1$. Bentuk matriks pada persamaan (2.2) adalah sebagai berikut:

$$\underline{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}; \quad \mathbf{X} = \begin{pmatrix} 1 & x_{11} & x_{21} & \cdots & x_{p1} \\ 1 & x_{12} & x_{22} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{pn} \end{pmatrix}; \quad \underline{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} \text{ dan } \underline{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Parameter $\underline{\beta}$ diestimasi dengan menggunakan metode *ordinary least square* (OLS) dengan meminimumkan $\underline{\varepsilon}'\underline{\varepsilon}$ yaitu dengan menurunkan $\underline{\varepsilon}'\underline{\varepsilon} = (\underline{y} - \mathbf{X}\underline{\beta})'(\underline{y} - \mathbf{X}\underline{\beta})$ terhadap $\underline{\beta}$. Hasil turunannya kemudian disamakan dengan nol sehingga diperoleh estimator $\hat{\underline{\beta}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\underline{y}$.

2.1.2. Model Regresi Nonparametrik

Pendekatan regresi parametrik tidak dapat digunakan ketika pola hubungan antara respon dengan prediktor tidak diketahui bentuk kurva regresinya, karena jika dipaksakan dan bentuk kurvanya tidak sesuai maka akan

menghasilkan varians *error* yang besar. Pendekatan yang sebaiknya digunakan pada kondisi ini adalah pendekatan regresi nonparametrik (Wu dan Zhang, 2006). Pendekatan regresi nonparametrik yang menyatakan hubungan antara p prediktor dengan respon tunggal yang melibatkan n pengamatan dapat dimodelkan dalam bentuk persamaan regresi sebagai berikut :

$$y_i = f(x_{1i}, x_{2i}, \dots, x_{pi}) + \varepsilon_i; i = 1, 2, \dots, n. \quad (2.3)$$

Keterangan:

y_i : respon pada pengamatan ke- i ,

x_{ji} : prediktor ke- j ($j = 1, 2, \dots, p$) pada pengamatan ke- i ,

f : fungsi/kurva regresi nonparametrik,

n : banyaknya pengamatan,

p : banyaknya prediktor,

ε_i : *random error* pada pengamatan ke- i .

Random error ε_i diasumsikan saling bebas berdistribusi normal dengan *mean* nol dan varians σ^2 .

Bentuk kurva regresi nonparametrik atau f dalam persamaan (2.3) tidak diketahui bentuknya. Bentuk kurva regresi f hanya diasumsikan mulus (*smooth*) yang termuat dalam suatu ruang fungsi tertentu, dimana pemilihan ruang fungsi berdasarkan sifat kemulusan pada kurva regresi f tersebut (Wahba, 1990). Pendekatan regresi nonparametrik lebih fleksibel dibandingkan dengan regresi parametrik dan diharapkan data sendiri yang mencari bentuk estimasinya (Eubank, 1999). Terdapat beberapa metode dalam regresi nonparametrik yang dapat digunakan untuk mengestimasi bentuk kurva regresi f , diantaranya adalah metode MARS yang menjadi fokus penelitian dalam disertasi ini. Metode MARS merupakan pengembangan dari metode *truncated* spline yang diuraikan pula dalam bab ini.

2.2. *Truncated* Spline dalam Regresi Nonparametrik

Spline merupakan salah satu metode yang cukup populer dalam pendekatan regresi nonparametrik. Spline sebagai pendekatan pola data

diperkenalkan oleh Whittaker pada tahun 1923, sedangkan spline sebagai persoalan optimasi dikembangkan oleh Reinsch pada tahun 1967 (Wahba, 1990). Metode spline dalam regresi nonparametrik dapat ditemui dalam banyak bentuk, diantaranya adalah *smoothing* spline (Wang, 1998; Kim *et al.*, 2004; Howell, 2007; Aydin, 2008) dan *truncated* spline (Smith, 1979; Budiantara, 2006; Sudiarsa *et al.*, 2015; Ruliana *et al.*, 2015). Kedua metode spline tersebut masing-masing menggunakan parameter yang berbeda untuk membuat estimasi kurva regresi lebih fleksibel, yaitu parameter penghalus pada *smoothing* spline dan titik knot pada *truncated* spline. Perbedaan jenis parameter tersebut, menyebabkan optimasi untuk mendapatkan estimator pada kedua metode spline tersebut juga berbeda. Estimator *smoothing* spline yang bergantung pada parameter penghalus diperoleh dengan optimasi *penalized least square* (PLS) (Wahba, 1990; Eubank, 1999), sedangkan estimator *truncated* spline yang bergantung pada titik knot diperoleh dengan optimasi OLS (Budiantara, 2006). Kajian tentang *truncated* spline selanjutnya akan lebih difokuskan pada sub bab ini, karena jenis metode spline ini merupakan dasar metode MARS yang akan dikembangkan dalam disertasi ini.

Pendekatan spline umumnya menyatakan hubungan antara p prediktor dengan respon tunggal pada persamaan (2.3), dalam model aditif berikut ini:

$$y_i = f_1(x_{1i}) + f_2(x_{2i}) + \dots + f_p(x_{pi}) + \varepsilon_i; i = 1, 2, \dots, n.$$

Persamaan ini dapat ditulis dalam bentuk:

$$y_i = \sum_{j=1}^p f_j(x_{ji}) + \varepsilon_i; i = 1, 2, \dots, n. \quad (2.4)$$

Apabila fungsi regresi f_j dalam persamaan (2.4) dihampiri dengan fungsi *truncated* spline yang memiliki q orde polinomial dan r knot, maka persamaan (2.4) dapat dinyatakan dalam bentuk sebagai berikut:

$$y_i = \theta_0 + \sum_{j=1}^p \left(\sum_{l=1}^q \theta_{jl} x_{ji}^l + \sum_{h=1}^r \gamma_{jh} (x_{ji} - t_{jh})_+^q \right) + \varepsilon_i; i = 1, 2, \dots, n, \quad (2.5)$$

dimana:

$$\theta_0 + \sum_{j=1}^p \left(\sum_{l=1}^q \theta_{jl} x_{ji}^l + \sum_{h=1}^r \gamma_{jh} (x_{ji} - t_{jh})_+^q \right) = \sum_{j=1}^p f_j(x_{ji}),$$

dengan :

$$(x_{ji} - t_{jh})_+^q = \begin{cases} (x_{ji} - t_{jh})^q & , \text{ jika } x_{ji} \geq t_{jh} \\ 0 & , \text{ jika } x_{ji} < t_{jh} \end{cases} . \quad (2.6)$$

Keterangan:

y_i : respon pada pengamatan ke- i ,

x_{ji} : prediktor ke- j pada pengamatan ke- i ,

θ_0 : intersep,

θ_{jl} : koefisien polinomial pada prediktor ke- j dan orde ke- l ,

γ_{jh} : koefisien *truncated* pada prediktor ke- j dan knot ke- h ,

t_{jh} : nilai knot pada prediktor ke- j dan knot ke- h ,

r : banyaknya knot,

q : orde polinomial *truncated* spline, jika $q = 1$ disebut *truncated* spline linier, $q = 2$ disebut *truncated* spline kuadratik dan $q = 3$ *truncated* spline kubik.

p : banyaknya prediktor,

n : banyaknya pengamatan,

ε_i : *random error* pada pengamatan ke- i , yang diasumsikan saling bebas berdistribusi normal dengan *mean* nol dan varians σ^2 .

Persamaan (2.5) untuk n data pengamatan dapat dinyatakan dalam bentuk matriks sebagai berikut :

$$\underline{y} = \mathbf{S}(\underline{x}, \underline{t})\underline{\theta} + \underline{\varepsilon}, \quad (2.7)$$

dengan:

$$\underline{y} = (y_1, y_2, \dots, y_n)',$$

$$\underline{\theta} = (\theta_0, \theta_{11}, \dots, \theta_{1q}, \gamma_{11}, \dots, \gamma_{1r}, \dots, \theta_{p1}, \dots, \theta_{pq}, \gamma_{p1}, \dots, \gamma_{pr})',$$

$$\underline{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)',$$

$$\mathbf{S}(\underline{x}, \underline{t}) = \begin{pmatrix} 1 & x_{11} & \dots & x_{11}^q & (x_{11} - t_{11})_+^q & \dots & (x_{11} - t_{1r})_+^q & \dots & x_{p1} & \dots & x_{p1}^q & (x_{p1} - t_{p1})_+^q & \dots & (x_{p1} - t_{pr})_+^q \\ 1 & x_{12} & \dots & x_{12}^q & (x_{12} - t_{11})_+^q & \dots & (x_{12} - t_{1r})_+^q & \dots & x_{p2} & \dots & x_{p2}^q & (x_{p2} - t_{p1})_+^q & \dots & (x_{p2} - t_{pr})_+^q \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \dots & x_{1n}^q & (x_{1n} - t_{11})_+^q & \dots & (x_{1n} - t_{1r})_+^q & \dots & x_{pn} & \dots & x_{pn}^q & (x_{pn} - t_{p1})_+^q & \dots & (x_{pn} - t_{pr})_+^q \end{pmatrix}.$$

Estimasi fungsi regresi *truncated* spline pada persamaan (2.7), diperoleh dengan menggunakan optimasi OLS, yaitu dengan menyelesaikan persamaan sebagai berikut :

$$\underset{\theta \in R^{1+p(q+r)}}{\text{Min}} (\varepsilon' \varepsilon) = \underset{\theta \in R^{1+p(q+r)}}{\text{Min}} \left((y - S(x, t)\theta)' (y - S(x, t)\theta) \right), \quad (2.8)$$

dengan penjabaran dan menurunkan $\varepsilon' \varepsilon$ terhadap θ dan menyamakan hasilnya dengan nol, maka diperoleh hasil estimator untuk θ dengan persamaan sebagai berikut:

$$\hat{\theta} = (S'(x, t)S(x, t))^{-1} S'(x, t)y. \quad (2.9)$$

Bentuk estimasi fungsi regresi *truncated* spline selanjutnya diperoleh berdasarkan persamaan (2.9) sebagai berikut :

$$\begin{aligned} \hat{f}(x) &= S(x, t)\hat{\theta}, \\ &= S(x, t)(S'(x, t)S(x, t))^{-1} S'(x, t)y, \\ &= A(\cdot)y, \end{aligned}$$

dengan:

$$A(\cdot) = S(x, t)(S'(x, t)S(x, t))^{-1} S'(x, t). \quad (2.10)$$

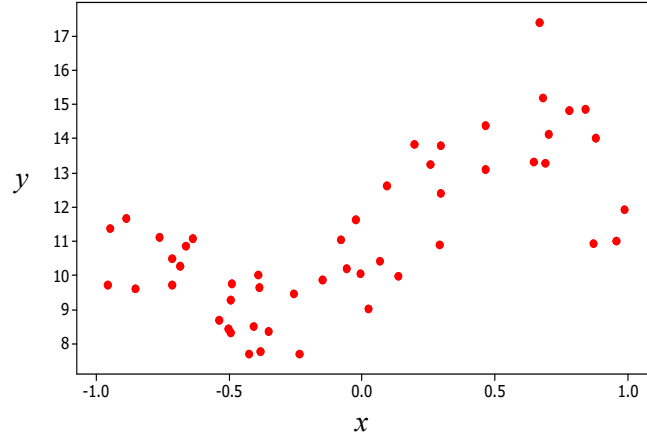
Matriks $A(\cdot)$ dalam persamaan (2.10) merupakan matriks *hat* yang bergantung pada titik-titik knot $t_{11}, \dots, t_{1r}, \dots, t_{p1}, \dots, t_{pr}$. Letak dan banyaknya titik knot tersebut harus dipilih secara optimal untuk mendapatkan estimator yang sesuai dengan data. Pemilihan titik knot optimal dapat diperoleh dengan berbagai metode, salah satunya dengan metode *GCV* yang dirumuskan dalam persamaan berikut ini:

$$\text{GCV}(t) = \frac{n^{-1} \left\| (\mathbf{I} - A(\cdot)) y \right\|^2}{\left[n^{-1} \text{tr}(\mathbf{I} - A(\cdot)) \right]^2}, \quad (2.11)$$

dengan norm $\|\cdot\|$ merupakan panjang vektor .

Titik-titik knot optimal $t_{11(\text{opt})}, \dots, t_{1r(\text{opt})}, \dots, t_{p1(\text{opt})}, \dots, t_{pr(\text{opt})}$ diperoleh dengan menentukan nilai titik knot yang meminimumkan nilai *GCV* pada persamaan (2.11).

Ilustrasi berikut ini memberikan contoh pemilihan titik knot optimal pada pendekatan *truncated* spline linier yang melibatkan satu prediktor, dengan contoh pola hubungan antar satu respon dengan satu prediktor diberikan oleh Gambar 2.1.

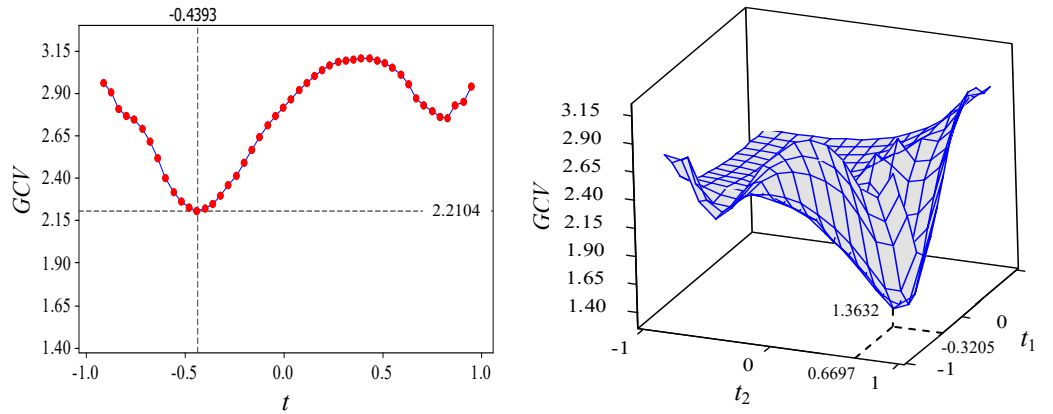


Gambar 2.1. Ilustrasi Pola Hubungan antara 1 Respon dengan 1 Prediktor

Pemilihan titik knot optimal pada Gambar 2.1 dapat dilakukan dengan tahapan sebagai berikut:

1. Menentukan banyaknya titik knot yang mungkin berdasarkan perubahan pola perilaku data yang terjadi pada sub-sub interval, seperti pada Gambar 2.1 tampak bahwa banyaknya titik knot yang mungkin adalah 1, 2, 3 atau 4 knot.
2. Menentukan kemungkinan letak titik knot (t) dalam interval data prediktor, yaitu dengan cara membagi range data x kedalam beberapa titik (misalnya z titik), sehingga diperoleh $x_{\min} < t_{[1]} < t_{[2]} < \dots < t_{[z]} < x_{\max}$. Contoh pada Gambar 2.1, kemungkinan letak titik knot berada pada: $-1 < t_{[1]} < t_{[2]} < \dots < t_{[z]} < 1$. Apabila banyaknya titik knot yang digunakan lebih dari satu, maka ditentukan pula minimum observasi diantara letak titik knot tersebut.
3. Mendapatkan nilai GCV minimum secara iteratif, untuk masing-masing banyaknya knot yang disebutkan pada tahap 1 menggunakan kemungkinan letak titik knot yang disebutkan pada tahap 2. Formula GCV diberikan oleh persamaan (2.11).
4. Membandingkan nilai GCV yang dihasilkan oleh kombinasi antara banyaknya titik knot dan letak titik knot yang disebutkan dalam tahap 3. Kombinasi knot yang menghasilkan nilai GCV terkecil ditetapkan sebagai knot optimal, yang selanjutnya digunakan dalam menyusun model.

Pola nilai GCV yang dihasilkan oleh masing-masing titik knot berdasarkan tahapan pemilihan titik knot optimal pada Gambar 2.1, dapat ditunjukkan dalam bentuk grafik untuk banyaknya satu titik knot dan dua titik knot, seperti disajikan pada Gambar 2.2.



(a) Nilai-Nilai GCV untuk 1 Titik Knot (b) Nilai-Nilai GCV untuk 2 Titik Knot

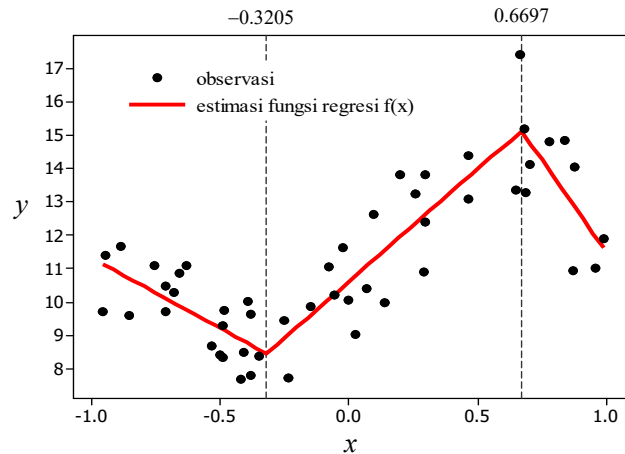
Gambar 2.2. Ilustrasi Pola Nilai GCV pada Berbagai Posisi Titik Knot dalam Estimator *Truncated Spline*

Gambar 2.2 menunjukkan bahwa titik knot optimal pada 1 titik knot berada pada nilai $t_{(opt)} = -0,4393$, sedangkan pada 2 titik knot berada pada nilai $t_{1(opt)} = -0,3205$ dan $t_{2(opt)} = 0,6697$. Hasil perbandingan nilai GCV yang dihasilkan oleh banyaknya titik knot selengkapnya disajikan dalam tabel 2.1.

Tabel 2.1. Perbandingan Nilai GCV berdasarkan Banyaknya Titik Knot dalam Estimator *Truncated Spline*

Banyaknya Titik Knot	Letak Titik Knot Optimal	Nilai GCV
1 knot	-0,4393	2,2104
2 knot	-0,3205, 0,6697	1,3632
3 knot	-0,6373, -0,4393, 0,7489	1,3643
4 knot	-0,6373, -0,4393, 0,8281, 0,9073	1,3882

Tabel 2.1 menunjukkan bahwa nilai GCV minimum yaitu 1,3632, dihasilkan oleh 2 titik knot yang berada pada titik $-0,3205$ dan $0,6697$ sehingga ditetapkan sebagai titik knot optimal. Letak dari titik knot optimal tersebut dapat ditunjukkan dalam Gambar 2.3.



Gambar 2.3. Ilustrasi Letak Titik Knot Optimal dari Estimasi Fungsi Regresi *Truncated Spline*

2.3. *Multivariate Adaptive Regression Spline (MARS)*

Metode *truncated spline* dalam penerapannya seringkali memiliki keterbatasan untuk menentukan posisi dan banyaknya titik knot yang digunakan ketika pemodelan regresi melibatkan prediktor yang banyak. Hal ini disebabkan karena kombinasi knot yang akan dipilih menjadi sangat banyak dan kompleks, yaitu dari kombinasi jumlah prediktor, posisi knot dan banyaknya knot (Hastie *et al.*, 2008). Metode MARS mampu mengatasi kelemahan *truncated spline* pada kasus ini, karena penentuan knot pada MARS tidak dicari secara satu persatu dari kombinasi tersebut tetapi dengan proses adaptif. Proses adaptif dalam MARS dilakukan dengan algoritma *stepwise* meliputi *forward* dan *backward stepwise*.

Terdapat beberapa penelitian yang membandingkan kemampuan metode MARS dengan metode lainnya, diantaranya adalah Abraham *et al.* (2001) yang mengkaji bahwa MARS menghasilkan komputasi yang lebih cepat dan tingkat akurasi yang lebih baik dibandingkan dengan metode NN pada pemodelan data *gas furnace*. Xiong *et al.* (2002) mengkaji bahwa MARS memberikan tingkat akurasi yang lebih baik dibandingkan dengan regresi logistik pada pemodelan

aktifitas air, pH dan Potassium pada pertumbuhan *Sacharomyces Cerevisiac*. Otok (2005) mengkaji bahwa MARS memberikan hasil yang lebih baik dibandingkan CART pada pengklasifikasian bank karena MARS menghasilkan model yang kontinu pada knot. Penelitian lain ada pula yang mengkombinasikan MARS dengan metode lainnya, seperti Lu *et al.* (2010) yang mengkaji *hybrid* MARS dengan SVR yang diterapkan pada peramalan penjualan produk teknologi informasi di China. Bakin *et al.* (1991) mengembangkan MARS menggunakan fungsi basis B-Spline.

Friedman (1991) menyatakan bahwa MARS merupakan kombinasi yang kompleks antara *truncated* spline dengan RPR. Hasil kombinasi tersebut menjadikan MARS memiliki beberapa kelebihan dibandingkan dengan *truncated* spline dan RPR. Beberapa kelebihan MARS dibandingkan dengan *truncated* spline adalah MARS mampu mengakomodir pengaruh aditif dan pengaruh interaksi antara prediktor dalam pemodelannya, sedangkan *truncated spline* hanya mengakomodir pengaruh aditif. MARS dapat digunakan pada pemodelan regresi yang melibatkan respon kontinu maupun kategori, sedangkan *truncated* spline umumnya digunakan pada respon kontinu saja. MARS juga memiliki kelebihan dari sisi waktu komputasi untuk pemodelan data yang melibatkan banyak prediktor dibandingkan dengan *truncated* spline, karena pemilihan knot pada MARS dilakukan dengan prosedur adaptif yang meliputi *forward* dan *backward stepwise*. MARS juga memiliki kelebihan dibandingkan dengan RPR, karena MARS menghasilkan model yang kontinu pada knot, yang tidak ditemui pada metode RPR (Friedman, 1991).

2.3.1. Fungsi MARS

Friedman (1991) menyatakan bahwa fungsi MARS diperoleh dengan algoritma *stepwise* (*forward* dan *backward*) dan dapat dinyatakan dalam bentuk persamaan sebagai berikut:

$$\begin{aligned} f(x_{1i}, x_{2i}, \dots, x_{pi}) &= \alpha_0 + \sum_{m=1}^M \alpha_m \prod_{k=1}^{K_m} [s_{km}(x_{v(k,m)i} - t_{km})]_+; \quad i = 1, 2, \dots, n, \quad (2.12) \\ &= \alpha_0 + \sum_{m=1}^M \alpha_m B_{mi}(x, t), \end{aligned}$$

dimana :

$$B_{mi}(\underline{x}, \underline{t}) = \prod_{k=1}^{K_m} [s_{km} (x_{v(k,m)i} - t_{km})]_+ ,$$

dengan : $x_{v(k,m)} \in \{x_j\}_{j=1}^p$ dan $t_{km} \in \{x_{v(k,m)i}\}_{i=1}^n$,

jika $s_{km} = +1$, maka $[+(x_{v(k,m)i} - t_{km})]_+ = \begin{cases} x_{v(k,m)i} - t_{km}, & \text{jika } x_{v(k,m)i} > t_{km}, \text{ dan} \\ 0 & , \text{ sebaliknya} \end{cases}$

jika $s_{km} = -1$, maka $[-(x_{v(k,m)i} - t_{km})]_+ = \begin{cases} t_{km} - x_{v(k,m)i}, & \text{jika } t_{km} > x_{v(k,m)i} \\ 0 & , \text{ sebaliknya} \end{cases}$.

Keterangan:

f : fungsi MARS,

α_0 : parameter fungsi basis konstan,

α_m : parameter fungsi basis *nonconstant* ke- m ,

$B_m(\underline{x}, \underline{t})$: fungsi basis ke- m ,

M : banyaknya fungsi basis *nonconstant*,

K_m : maksimum interaksi pada fungsi basis ke- m ,

s_{km} : tanda fungsi basis pada interaksi ke- k dan fungsi basis ke- m , s_{km} bernilai +1 atau -1,

$x_{v(k,m)}$: variabel x ke- v , dimana v adalah indeks variabel x yang berkaitan dengan interaksi ke- k dan fungsi basis ke- m dalam fungsi MARS,

t_{km} : nilai knot pada interaksi ke- k dan fungsi basis ke- m .

Fungsi MARS pada persamaan (2.12) dapat didekomposisikan menjadi beberapa fungsi berdasarkan variabel yang masuk dalam model, baik yang memuat satu variabel maupun interaksi antara variabel. Misalkan $f_a(x_{ai})$ adalah fungsi yang memuat satu variabel, $f_{ab}(x_{ai}, x_{bi})$ adalah fungsi yang memuat interaksi antara dua variabel, $f_{abc}(x_{ai}, x_{bi}, x_{ci})$ adalah fungsi yang memuat interaksi antara tiga variabel, dengan $x_{ai}, x_{bi}, x_{ci} \in \{x_{ji}\}_{j=1}^p$, maka persamaan (2.12) dapat didekomposisikan menjadi:

$$f(x_{1i}, x_{2i}, \dots, x_{pi}) = \alpha_0 + \sum_{K_m=1} f_a(x_{ai}) + \sum_{K_m=2} f_{ab}(x_{ai}, x_{bi}) + \sum_{K_m=3} f_{abc}(x_{ai}, x_{bi}, x_{ci}) \\ + \dots + \sum_{K_m=p} f_{123\dots p}(x_{1i}, x_{2i}, \dots, x_{pi}); \quad i = 1, 2, \dots, n. \quad (2.13)$$

Persamaan (2.13) menunjukkan bahwa penjumlahan pertama meliputi semua fungsi untuk satu variabel, penjumlahan kedua meliputi semua fungsi untuk interaksi antara dua variabel, penjumlahan ketiga meliputi semua fungsi untuk interaksi antara tiga variabel dan seterusnya hingga meliputi semua fungsi untuk interaksi antara p variabel.

Misalkan $V(m) = \{v(k, m)\}_{k=1}^{K_m}$ adalah himpunan dari indeks variabel yang termuat dalam fungsi basis ke- m (B_m), maka setiap fungsi univariat pada penjumlahan pertama dalam persamaan (2.13) dapat dinyatakan sebagai:

$$f_a(x_{ai}) = \sum_{\substack{K_m=1 \\ a \in V'(m)}} \alpha_m B_m(x_{ai}, t_a). \quad (2.14)$$

Fungsi $f_a(x_{ai})$ merupakan penjumlahan semua fungsi basis yang hanya meliputi variabel x_{ai} , yang menyatakan fungsi dari satu variabel saja. Setiap fungsi bivariat pada persamaan (2.13) dinyatakan sebagai:

$$f_{ab}(x_{ai}, x_{bi}) = \sum_{\substack{K_m=2 \\ (a,b) \in V'(m)}} \alpha_m B_m[(x_{ai}, t_a), (x_{bi}, t_b)]. \quad (2.15)$$

Fungsi $f_{ab}(x_{ai}, x_{bi})$ merupakan penjumlahan semua fungsi basis untuk interaksi antara dua variabel x_{ai} dan x_{bi} . Penambahan fungsi (2.15) dengan fungsi univariat pada persamaan (2.14) menghasilkan:

$$f_{ab}^*(x_{ai}, x_{bi}) = f_a(x_{ai}) + f_b(x_{bi}) + f_{ab}(x_{ai}, x_{bi}). \quad (2.16)$$

Persamaan (2.16) menyatakan fungsi dari dua variabel x_{ai} dan x_{bi} secara bersama-sama. Setiap fungsi trivariat pada penjumlahan ketiga dalam persamaan (2.13) juga dapat dinyatakan sebagai:

$$f_{abc}(x_{ai}, x_{bi}, x_{ci}) = \sum_{\substack{K_m=3 \\ (a,b,c) \in V'(m)}} \alpha_m B_m[(x_{ai}, t_a), (x_{bi}, t_b), (x_{ci}, t_c)]. \quad (2.17)$$

Fungsi $f_{abc}(x_{ai}, x_{bi}, x_{ci})$ merupakan penjumlahan semua fungsi basis untuk interaksi antara tiga variabel x_{ai} , x_{bi} dan x_{ci} . Penambahan fungsi (2.17) dengan fungsi univariat dan bivariat pada persamaan (2.14) dan (2.15) menghasilkan:

$$f_{abc}^*(x_{ai}, x_{bi}, x_{ci}) = f_a(x_{ai}) + f_b(x_{bi}) + f_c(x_{ci}) + f_{ab}(x_{ai}, x_{bi}) + f_{ac}(x_{ai}, x_{ci}) + f_{bc}(x_{bi}, x_{ci}) + f_{abc}(x_{ai}, x_{bi}, x_{ci}) \quad (2.18)$$

Persamaan (2.18) menyatakan fungsi dari tiga variabel x_{ai} , x_{bi} dan x_{ci} secara bersama-sama. Uraian berikutnya hingga dicapai fungsi dengan p interaksi, dilakukan dengan cara yang sama seperti sebelumnya.

2.3.2. Estimasi Fungsi Regresi MARS pada Respon Kontinu

Metode MARS dalam penerapannya dapat dibedakan berdasarkan jenis respon yang digunakan, yaitu MARS respon kontinu dan MARS respon kategorik. MARS respon kontinu jika responnya bertipe kuantitatif yaitu merupakan data interval atau rasio. MARS respon kategorik jika responnya bertipe kualitatif yaitu merupakan data nominal atau ordinal. Persamaan regresi MARS yang menyatakan hubungan antara p prediktor dengan respon tunggal kontinu yang melibatkan n pengamatan dapat diperoleh berdasarkan persamaan (2.12), sebagai berikut:

$$\begin{aligned} y_i &= f(x_{1i}, x_{2i}, \dots, x_{pi}) + \varepsilon_i; \quad i = 1, 2, \dots, n, \\ &= \alpha_0 + \sum_{m=1}^M \alpha_m \prod_{k=1}^{K_m} [s_{km}(x_{v(k,m)i} - t_{km})]_+ + \varepsilon_i, \\ &= \alpha_0 + \sum_{m=1}^M \alpha_m B_{mi}(x, t) + \varepsilon_i. \end{aligned} \quad (2.19)$$

Persamaan (2.19) dapat dinyatakan dalam bentuk matriks sebagai berikut :

$$\tilde{y}^* = \mathbf{B}^*(\tilde{x}, \tilde{t}) \tilde{\alpha}^* + \tilde{\varepsilon}^*, \quad (2.20)$$

dengan:

$$\tilde{y}^* = (y_1, y_2, \dots, y_n)', \quad \tilde{\alpha}^* = (\alpha_0, \alpha_1, \dots, \alpha_M)', \quad \tilde{\varepsilon}^* = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)' \text{ dan}$$

$$\mathbf{B}^*(\tilde{x}, \tilde{t}) = \begin{pmatrix} 1 & \prod_{k=1}^{K_1} [s_{k1}(x_{v(k,1)1} - t_{k1})]_+ & \dots & \prod_{k=1}^{K_M} [s_{kM}(x_{v(k,M)1} - t_{kM})]_+ \\ 1 & \prod_{k=1}^{K_1} [s_{k1}(x_{v(k,1)2} - t_{k1})]_+ & \dots & \prod_{k=1}^{K_M} [s_{kM}(x_{v(k,M)2} - t_{kM})]_+ \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \prod_{k=1}^{K_1} [s_{k1}(x_{v(k,1)n} - t_{k1})]_+ & \dots & \prod_{k=1}^{K_M} [s_{kM}(x_{v(k,M)n} - t_{kM})]_+ \end{pmatrix}. \quad (2.21)$$

Vektor \underline{y}^* dalam persamaan (2.20) merupakan vektor respon berukuran $n \times 1$ dan $\mathbf{B}^*(\underline{x}, \underline{t})$ merupakan matriks berorde $n \times (M+1)$. $\underline{\alpha}^*$ adalah vektor yang memuat koefisien regresi berorde $(M+1) \times 1$. $\underline{\varepsilon}^*$ adalah vektor *random error* berorde $n \times 1$ yang diasumsikan saling bebas berdistribusi normal dengan $E(\underline{\varepsilon}^*) = \underline{0}$ dan $Var(\underline{\varepsilon}^*) = \sigma^2 \mathbf{I}$.

Estimasi parameter $\underline{\alpha}^*$ dalam persamaan (2.20) diperoleh dengan menyelesaikan optimasi OLS yaitu dengan meminimumkan jumlah kuadrat *error* sebagai berikut:

$$\underset{\underline{\alpha} \in R^{M+1}}{\text{Min}} \left(\underline{\varepsilon}^{*'} \underline{\varepsilon}^* \right) = \underset{\underline{\alpha} \in R^{M+1}}{\text{Min}} \left(\left(\underline{y}^* - \mathbf{B}^*(\underline{x}, \underline{t}) \underline{\alpha}^* \right)' \left(\underline{y}^* - \mathbf{B}^*(\underline{x}, \underline{t}) \underline{\alpha}^* \right) \right). \quad (2.22)$$

Penyelesaian optimasi dari persamaan (2.22) diperoleh dengan menurunkan secara parsial hasil penjabaran $\underline{\varepsilon}^{*'} \underline{\varepsilon}^*$ terhadap $\underline{\alpha}$ dan menyamakan hasilnya dengan nol, sehingga diperoleh estimasi untuk $\underline{\alpha}^*$ sebagai berikut:

$$\hat{\underline{\alpha}}^* = \left(\mathbf{B}^{*'}(\underline{x}, \underline{t}) \mathbf{B}^*(\underline{x}, \underline{t}) \right)^{-1} \mathbf{B}^{*'}(\underline{x}, \underline{t}) \underline{y}^*. \quad (2.23)$$

Bentuk estimasi fungsi regresi MARS respon kontinu, selanjutnya dapat diperoleh berdasarkan persamaan (2.23) dengan persamaan sebagai berikut:

$$\begin{aligned} \hat{f}^*(\underline{x}) &= \mathbf{B}^*(\underline{x}, \underline{t}) \hat{\underline{\alpha}}^*, \\ &= \mathbf{B}^*(\underline{x}, \underline{t}) \left(\mathbf{B}^{*'}(\underline{x}, \underline{t}) \mathbf{B}^*(\underline{x}, \underline{t}) \right)^{-1} \mathbf{B}^{*'}(\underline{x}, \underline{t}) \underline{y}^*, \\ &= \mathbf{H}^*(\cdot) \underline{y}^*, \end{aligned} \quad (2.24)$$

dengan:

$$\mathbf{H}^*(\cdot) = \mathbf{B}^*(\underline{x}, \underline{t}) \left(\mathbf{B}^{*'}(\underline{x}, \underline{t}) \mathbf{B}^*(\underline{x}, \underline{t}) \right)^{-1} \mathbf{B}^{*'}(\underline{x}, \underline{t}). \quad (2.25)$$

Matriks $\mathbf{H}^*(\cdot)$ dalam persamaan (2.25) merupakan matriks *hat* yang sangat bergantung pada matriks $\mathbf{B}^*(\underline{x}, \underline{t})$ yang memuat fungsi basis $B_1(\underline{x}, \underline{t}), B_2(\underline{x}, \underline{t}), \dots, B_M(\underline{x}, \underline{t})$. Fungsi basis tersebut harus dipilih secara optimal untuk mendapatkan estimator MARS yang sesuai dengan data. Pemilihan fungsi basis optimal dalam metode MARS dilakukan dengan prosedur *stepwise (forward dan backward)* berdasarkan nilai *GCV* minimum.

2.3.3. Pemilihan Fungsi Basis Optimal dalam MARS

Pemilihan fungsi basis optimal dalam metode MARS dilakukan dengan prosedur *stepwise* yang meliputi *forward stepwise* dan *backward stepwise*. *Forward stepwise* membangun model MARS dengan menambahkan fungsi basis *truncated* spline (knot dan interaksi) hingga diperoleh model dengan jumlah fungsi basis maksimum. *Backward stepwise* dilakukan setelah proses *forward stepwise* selesai, yang bertujuan untuk menentukan jumlah fungsi basis yang layak dalam model. Tahapan ini menghapus fungsi basis yang kontribusinya paling kecil terhadap nilai dugaan respon berdasarkan nilai *GCV* minimum.

a. *Forward Stepwise* :

Menurut Friedman (1991), prosedur *forward stepwise* dimulai dengan menetapkan fungsi basis konstan yaitu $B_0 = 1$. Misalkan terdapat $M + 1$ fungsi basis yaitu $B_0, B_1(\underline{x}, \underline{t}), B_2(\underline{x}, \underline{t}), \dots, B_M(\underline{x}, \underline{t})$, maka setiap tahap penambahan fungsi basis dalam *forward stepwise* dilakukan dengan memasukkan pasangan fungsi basis baru sebagai berikut:

$$\begin{aligned} B_{M+1}(\underline{x}, \underline{t}) &= B_M(\underline{x}, \underline{t}) \left[+ (x_{v(k,m)} - t_{km}) \right]_+, \text{ dan} \\ B_{M+2}(\underline{x}, \underline{t}) &= B_M(\underline{x}, \underline{t}) \left[- (x_{v(k,m)} - t_{km}) \right]_+. \end{aligned} \quad (2.26)$$

Fungsi basis $B_m(\underline{x}, \underline{t})$ pada persamaan (2.26) adalah fungsi basis induk (*parent*) yang merupakan anggota dari himpunan fungsi basis yang telah ada sebelum dilakukan penambahan pasangan fungsi basis baru, pada contoh kasus diatas yaitu: $B_m(\underline{x}, \underline{t}) \in \{B_0, B_1(\underline{x}, \underline{t}), B_2(\underline{x}, \underline{t}), \dots, B_M(\underline{x}, \underline{t})\}$. Persamaan (2.26) menspesifikasikan $x_{v(k,m)}$ adalah prediktor yang tidak terdapat dalam fungsi basis *parent* $B_m(\underline{x}, \underline{t})$ dan $t_{km} \in \{x_{v(k,m)i}\}_{i=1,2,\dots,n}$, untuk $x_{v(k,m)} \in \{x_j\}_{j=1,\dots,p}$. Nilai n adalah banyaknya pengamatan sedangkan p adalah banyaknya prediktor.

Penambahan pasangan fungsi basis baru pada *forward stepwise* dilakukan hingga dicapai jumlah fungsi basis maksimum (M_{\max}). Pemilihan pasangan fungsi basis baru yang ditambahkan, didasarkan pada

nilai err terkecil (Jekabsons, 2011). Nilai err diperoleh dari nilai MSE pada setiap penambahan pasangan fungsi basis dibagi dengan nilai varians respon, dengan persamaan sebagai berikut:

$$err(B) = \frac{\sum_{i=1}^n (y_i - \hat{y}(B))^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.27)$$

dimana $err(B)$ adalah nilai err model pada setiap penambahan pasangan fungsi basis baru, y_i adalah respon pada pengamatan ke- i , $\hat{y}(B)$ adalah prediksi respon pada setiap penambahan pasangan fungsi basis baru dan \bar{y} adalah nilai rata-rata respon.

b. Backward Stepwise :

Prosedur *backward stepwise* dilakukan setelah diperoleh jumlah fungsi basis maksimum (M_{\max}) yang dihasilkan oleh prosedur *forward stepwise*. Prosedur ini dimulai dengan menyeleksi satu persatu fungsi basis *forward stepwise* kecuali $B_0 = 1$ (fungsi basis konstan) dan mengeluarkannya jika nilai GCV menurun ketika fungsi basis itu dibuang. Proses penghapusan fungsi basis dilanjutkan sampai nilai GCV tidak turun meskipun fungsi basis yang masih tersisa dibuang. Fungsi basis yang tersisa dari prosedur *backward stepwise* ini adalah fungsi basis optimal.

Friedman (1991) menggunakan metode GCV dalam prosedur *backward stepwise* untuk memilih fungsi basis optimal pada algoritma MARS. Metode GCV yang digunakan dalam MARS dikembangkan dari Craven dan Wahba (1979) menggunakan fungsi kompleks dari Friedman dan Silverman (1989). Formula GCV ini dinyatakan sebagai berikut:

$$GCV(B) = \frac{MSE}{\left[1 - \frac{C(\tilde{M})}{n}\right]^2} = \frac{n^{-1} \left\| (\mathbf{I} - \mathbf{H}^*(\cdot)) \tilde{\mathbf{y}} \right\|^2}{\left[1 - \frac{C(\tilde{M})}{n}\right]^2}, \quad (2.28)$$

dengan: $C(\tilde{M}) = C(M) + \frac{d.M}{2}$.

Keterangan:

\underline{y} : vektor respon,

\mathbf{I} : matriks identitas,

$\mathbf{H}^*(\cdot)$: matriks *hat* dalam persamaan (2.25) yang berisi fungsi basis,

MSE : *mean square error*,

$C(\tilde{M})$: fungsi kompleks,

M : banyaknya fungsi basis *nonconstant*,

$C(M)$: banyaknya fungsi basis konstan dan *nonconstant* $(M + 1)$,

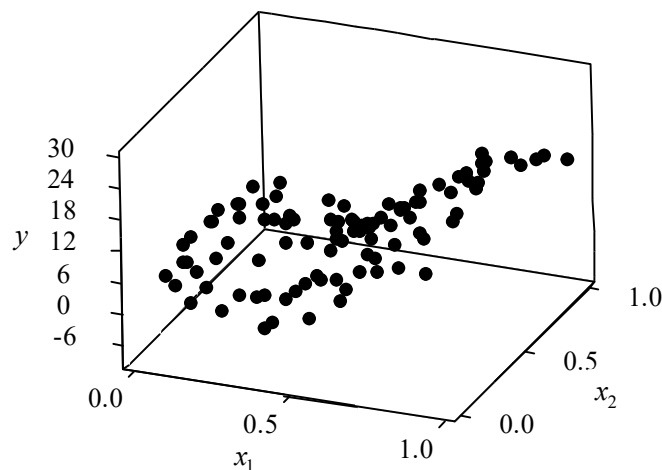
d : derajat interaksi, nilai $d = 2$ untuk model aditif dan $d = 3$ untuk model interaksi (Friedman, 1991; Jekabsons, 2011),

n : banyaknya pengamatan.

Ilustrasi berikut ini memberikan contoh pemilihan fungsi basis optimal berdasarkan tahapan *forward* dan *backward stepwise* dalam metode MARS. Misalkan diberikan persamaan regresi $y_i = f(x_{1i}, \dots, x_{pi}) + \varepsilon_i$, dimana fungsi regresi f mengikuti bentuk fungsi MARS sebagai berikut:

$$f(x_{1i}, x_{2i}) = 4 + 39(x_{1i} - 0,35)_+ + 14(0,35 - x_{1i})_+ - 17(x_{2i} - 0,5)_+ - 10(0,5 - x_{2i})_+ - 15(x_{1i} - 0,35)_+(x_{2i} - 0,5)_+ - 26(x_{1i} - 0,35)_+(0,5 - x_{2i})_+, \quad (2.29)$$

dengan $x_{1i}, x_{2i} \sim U(0,1)$, $\varepsilon_i \sim N(0,1)$ dan $i = 1, 2, \dots, 100$, sehingga diperoleh pola hubungan antara respon dan prediktor seperti pada Gambar 2.4.



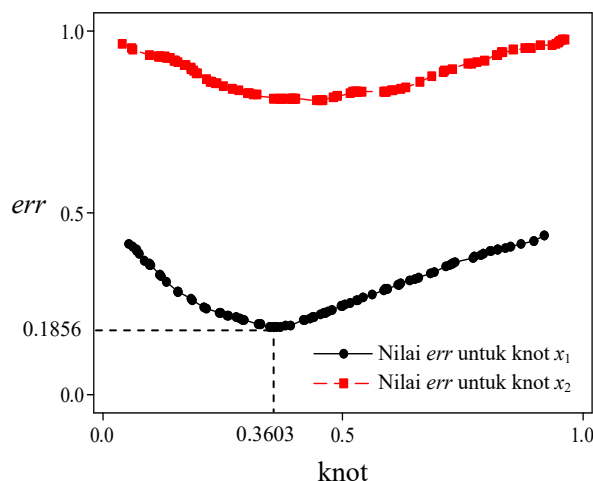
Gambar 2.4. Ilustrasi Pola Hubungan antara 1 Respon dan 2 Prediktor

Pada *forward stepwise*, fungsi basis yang diperoleh dari pola hubungan antara respon dan prediktor dalam Gambar 2.4 diperoleh berdasarkan nilai *err* (persamaan (2.27)) terkecil yang dihasilkan oleh setiap tahapan penambahan pasangan fungsi basis dalam model MARS, seperti yang disajikan dalam Tabel 2.2.

Tabel 2.2. Penambahan Fungsi Basis pada *Forward Stepwise* dalam Estimator MARS

Tahapan Penambahan Fungsi Basis	Fungsi Basis yang Ditambahkan	Nilai <i>err</i>
Fungsi Basis Konstan	$B_0 = 1$	1,0000
Penambahan ke-1	$B_1 = (x_1 - 0,3603)_+$	0,1856
	$B_2 = (0,3603 - x_1)_+$	
Penambahan ke-2	$B_3 = (x_2 - 0,4792)_+$	0,0306
	$B_4 = (0,4792 - x_2)_+$	
Penambahan ke-3	$B_5 = B_1(x_2 - 0,4564)_+$	0,0191
	$B_6 = B_1(0,4564 - x_2)_+$	
Penambahan ke-4	$B_7 = B_1(x_2 - 0,3060)_+$	0,0180
	$B_8 = B_1(0,3060 - x_2)_+$	
Penambahan ke-5	$B_9 = B_2(x_2 - 0,3943)_+$	0,0170
	$B_{10} = B_2(0,3943 - x_2)_+$	

Forward stepwise dalam model MARS selalu diawali dengan memasukkan fungsi basis konstan yaitu $B_0 = 1$ (intersep). Tabel 2.2 menunjukkan bahwa pemasukan fungsi basis konstan B_0 dalam tahapan ini menghasilkan nilai *err* dalam persamaan (2.27) sama dengan 1,0000. Penambahan berikutnya dilakukan dengan memasukkan pasangan fungsi basis ke-1 (B_1 dan B_2), yaitu: $(x_1 - 0,3603)_+$ dan $(0,3603 - x_1)_+$ dengan fungsi basis induk (*parent*) $B_0 = 1$. Penambahan pasangan fungsi basis yang mengandung prediktor x_1 dan knot $t = 0,3603$ ini dipilih, karena menurunkan nilai *err* model lebih besar yaitu menjadi 0,1856, dibandingkan dengan ketika pasangan fungsi basis lainnya (baik yang mengandung prediktor x_2 atau titik knot lainnya) yang ditambahkan, seperti ditunjukkan pada Gambar 2.5.



Gambar 2.5. Ilustrasi Pola Nilai *err* pada Penambahan Pasangan Fungsi Basis ke-1 (B_1 dan B_2) dalam Estimator MARS

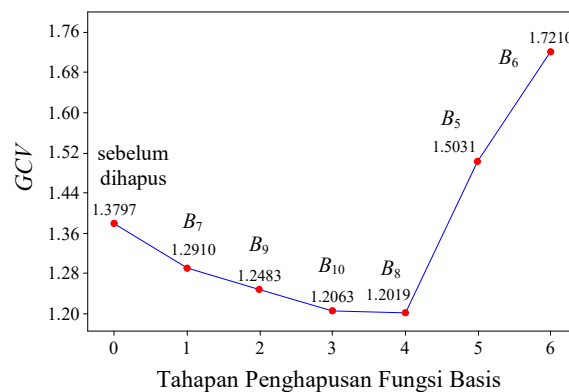
Penambahan pasangan fungsi basis berikutnya (dengan interaksi atau tanpa interaksi) terus dilakukan secara iteratif dengan cara yang sama seperti pada penambahan pasangan fungsi basis ke-1 dan berhenti ketika dicapai jumlah maksimum fungsi basis. Pada *forward stepwise*, jumlah maksimum fungsi basis dicapai/dibatasi ketika selisih nilai *err* antara setelah dengan sebelum penambahan pasangan fungsi basis bernilai kecil (Jekabsons, 2011), misalnya bernilai kurang dari 0,001. Pembatasan penambahan pasangan fungsi basis berdasarkan selisih nilai *err* tersebut, dimaksudkan untuk menghindari model yang sangat kompleks dengan tidak memasukkan pasangan fungsi basis yang kontribusinya sangat kecil terhadap model. Pada contoh ilustrasi dalam Tabel 2.2, penambahan pasangan fungsi basis berhenti pada penambahan ke-5 (B_9 dan B_{10}) dengan nilai *err* = 0,0170, sehingga jumlah maksimum fungsi basis pada akhir tahapan ini (*forward stepwise*) sama dengan 11 fungsi basis.

Backward stepwise dilakukan setelah penambahan fungsi basis dalam *forward stepwise* berhenti. *Backward stepwise* menghapus fungsi basis *forward stepwise* dalam Tabel 2.2 secara satu persatu berdasarkan nilai *GCV* (persamaan (2.28)) minimum. Penghapusan fungsi basis dalam *backward stepwise* terus dilakukan hingga tersisa fungsi basis konstan saja ($B_0 = 1$), seperti disajikan dalam Tabel 2.3.

Tabel 2.3. Penghapusan Fungsi Basis pada *Backward Stepwise* dalam Estimator MARS

Tahapan Penghapusan Fungsi Basis	Fungsi Basis yang Dihapus	Nilai <i>GCV</i>
Sebelum Dihapus	-	1,3797
Penghapusan ke-1	$B_7 = B_1(x_2 - 0,3060)_+$	1,2910
Penghapusan ke-2	$B_9 = B_2(x_2 - 0,3943)_+$	1,2483
Penghapusan ke-3	$B_{10} = B_2(0,3943 - x_2)_+$	1,2063
Penghapusan ke-4	$B_8 = B_1(0,3060 - x_2)_+$	1,2019
Penghapusan ke-5	$B_5 = B_1(x_2 - 0,4564)_+$	1,5031
Penghapusan ke-6	$B_6 = B_1(0,4564 - x_2)_+$	1,7210
Penghapusan ke-7	$B_2 = (0,3603 - x_1)_+$	3,0672
Penghapusan ke-8	$B_4 = (0,4792 - x_2)_+$	7,2777
Penghapusan ke-9	$B_3 = (x_2 - 0,4792)_+$	11,3510
Penghapusan ke-10	$B_1 = (x_1 - 0,3603)_+$	45,3820

Pola perubahan nilai *GCV* yang dihasilkan oleh penghapusan fungsi basis dalam Tabel 2.3, ditunjukkan dengan jelas dalam Gambar 2.6 yang diwakili oleh penghapusan ke-1 sampai dengan ke-6.



Gambar 2.6. Ilustrasi Pola Nilai *GCV* pada Penghapusan Fungsi Basis dalam *Forward Stepwise* dalam Estimator MARS

Tabel 2.3 menunjukkan bahwa penghapusan pertama dilakukan pada fungsi basis B_7 . Hal ini disebabkan karena pada penghapusan fungsi basis ini menyebabkan penurunan nilai *GCV* model yang lebih besar yaitu menjadi 1,2910, dibandingkan dengan ketika salah satu dari fungsi basis lainnya (B_0 ,

$B_1, B_2, B_3, B_4, B_5, B_6, B_8, B_9, B_{10}$) yang dihapus. Penghapusan ini mengurangi fungsi basis yang semula berjumlah 11 menjadi berjumlah 10. Penghapusan ke-2 hingga seterusnya, terus dilakukan dengan cara yang sama seperti pada penghapusan ke-1, hingga pada akhir penghapusan fungsi basis hanya satu fungsi basis saja yang tersisa yaitu fungsi basis konstan ($B_0 = 1$).

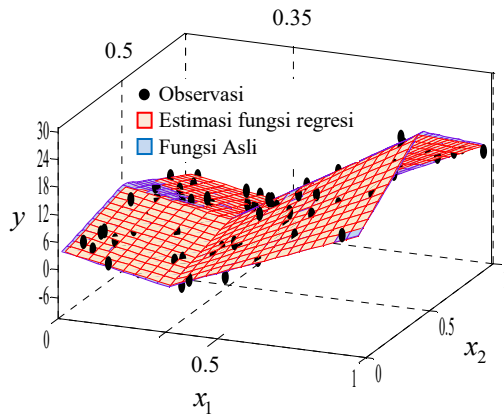
Tabel 2.3 dan Gambar 2.6 menunjukkan bahwa pola nilai GCV yang dihasilkan oleh tahapan penghapusan ini terus menurun pada penghapusan ke-1 sampai dengan ke-4 yang berturut-turut menghapus fungsi basis B_7, B_9, B_{10} dan B_8 . Hal sebaliknya terjadi pada penghapusan ke-5 sampai dengan ke-10 yang menghapus fungsi basis B_5, B_6, B_2, B_4, B_3 dan B_1 , dimana pola nilai GCV tidak turun tetapi terus meningkat seiring dengan penghapusan fungsi basis. Hal ini berarti bahwa fungsi basis yang harus dihapus pada tahapan *backward stepwise* ini adalah fungsi basis B_7, B_8, B_9 dan B_{10} , sedangkan fungsi basis B_1, B_2, B_3, B_4, B_5 dan B_6 tidak dihapus tetapi tetap dipertahankan dan dinyatakan sebagai fungsi basis optimal dalam metode MARS. Selanjutnya berdasarkan fungsi basis optimal tersebut, dapat diperoleh estimasi fungsi regresi MARS menggunakan persamaan (2.24), dengan hasil estimasi sebagai berikut:

$$\begin{aligned}\hat{f}(x_1, x_2) = & 4,29 + 38,87B_1 + 12,39B_2 - 14,94B_3 - 9,88B_4 \\ & - 16,21B_5 - 29,53B_6,\end{aligned}\tag{2.30}$$

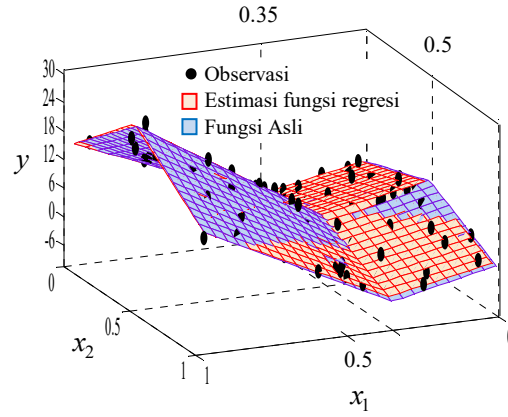
dengan :

$$\begin{aligned}B_1 &= (x_1 - 0,3603)_+, \\ B_2 &= (0,3603 - x_1)_+, \\ B_3 &= (x_2 - 0,4792)_+, \\ B_4 &= (0,4792 - x_2)_+, \\ B_5 &= B_1(x_2 - 0,4564)_+, \\ B_6 &= B_1(0,4564 - x_2)_+.\end{aligned}$$

Bentuk estimasi fungsi regresi dari persamaan (2.30), dapat disajikan dalam Gambar 2.7.



(a) Tampak dari Depan



(b) Tampak dari Belakang

Gambar 2.7. Ilustrasi Grafik Estimasi Fungsi Regresi MARS

2.4. Model Regresi Nonparametrik Birespon kontinu

Model regresi nonparametrik birespon kontinu merupakan salah satu model dalam regresi nonparametrik yang melibatkan dua respon kontinu, dan antara respon saling berkorelasi. Wang *et al.* (2000) dalam penelitiannya pada *smoothing spline*, menyatakan model regresi nonparametrik birespon kontinu untuk satu prediktor dalam bentuk persamaan regresi sebagai berikut:

$$y_{\ell i} = f_{\ell}(x_i) + \varepsilon_{\ell i}; \quad \ell = 1, 2; \quad i = 1, 2, \dots, n. \quad (2.31)$$

Keterangan :

$y_{\ell i}$: respon ke- ℓ pada pengamatan ke- i ,

f_{ℓ} : kurva regresi nonparametrik birespon pada respon ke- ℓ ,

x_i : prediktor pada pengamatan ke- i ,

n : banyaknya pengamatan,

ℓ : banyaknya respon,

$\varepsilon_{\ell i}$: *random error* pada respon ke ℓ dan pengamatan ke i .

Model regresi nonparametrik birespon kontinu dalam persamaan (2.31) mengasumsikan random error $\varepsilon_{\ell i}$ sebagai berikut:

$$E(\varepsilon_{\ell i}) = 0, \quad \text{cov}(\varepsilon_{\ell i}, \varepsilon_{\ell i'}) = \begin{cases} \sigma_{\ell\ell}, & \text{jika } i = i' \\ 0, & \text{jika } i \neq i' \end{cases}, \quad \text{dengan } \ell = 1, 2; \quad i = 1, 2, \dots, n,$$

dan:

$$\text{cov}(\varepsilon_{1i}, \varepsilon_{2i'}) = \begin{cases} \sigma_{12}, & \text{jika } i = i' \\ 0, & \text{jika } i \neq i' \end{cases} \quad (2.32)$$

Terdapat beberapa peneliti lain yang mengembangkan asumsi tersebut pada berbagai pola data dan pada metode regresi nonparametrik lainnya, seperti Lestari *et al.* (2012) mengembangkannya pada *smoothing spline* untuk kasus heteroskedastisitas. Wibowo *et al.* (2012) yang mengembangkan multirespon *smoothing spline* pada regresi semiparametrik. Fernandez *et al.* (2014) yang mengembangkan birespon *smoothing spline* pada data longitudinal. Peneliti lainnya yaitu Welsh *et al.* (2006) dan Chamidah *et al.* (2012) juga mengembangkan model tersebut masing-masing menggunakan metode lokal linier dan polinomial lokal.

Para peneliti umumnya mengembangkan model birespon/multirespon dengan menggunakan model aditif antar prediktor dalam pemodelan data. Para peneliti belum ada yang mengembangkan asumsi model birespon tersebut pada metode MARS, dimana metode MARS memiliki kelebihan yang tidak ditemui pada metode regresi nonparametrik lainnya. Kelebihan metode MARS diantaranya mampu mengakomodir pengaruh aditif dan pengaruh interaksi antar prediktor dalam pemodelannya, sehingga memungkinkan prediksi respon yang lebih baik.

Keberadaan korelasi respon pada model regresi birespon dapat diketahui dengan menggunakan koefisien korelasi Pearson yang diperoleh dengan persamaan sebagai berikut:

$$\rho(y_1, y_2) = \frac{\text{cov}(\tilde{y}_1, \tilde{y}_2)}{[\text{var}(\tilde{y}_1) \text{var}(\tilde{y}_2)]^{1/2}} \quad (2.33)$$

Nilai koefisien korelasi pada persamaan (2.33) berkisar antara -1 hingga 1 ($-1 \leq \rho(\tilde{y}_1, \tilde{y}_2) \leq 1$). Jika koefisien korelasi mendekati -1 atau 1 maka hubungan atau korelasi diantara respon semakin kuat, dan sebaliknya jika mendekati nol maka hubungan atau korelasi diantara respon semakin lemah (Draper dan Smith, 1998).

2.5. Ukuran Kesesuaian Model

Model yang sesuai adalah model yang memiliki ukuran yang baik pada tahap evaluasi. Dalam penelitian ini menggunakan dua ukuran yaitu *Root of Mean Square Error (RMSE)* serta R^2 atau koefisien determinasi. *RMSE* adalah ukuran yang digunakan untuk mengevaluasi model dengan cara menghitung selisih antara nilai observasi $y_{\ell i}$ dengan nilai estimasi $\hat{f}_{\ell i}$ pada setiap respon ($\ell = 1, 2$), dengan persamaan sebagai berikut (Greene, 2007):

$$RMSE_{y_{\ell}} = \sqrt{\frac{\sum_{i=1}^n (y_{\ell i} - \hat{f}_{\ell i})^2}{n-1}}, \ell = 1, 2. \quad (2.34)$$

Nilai R^2 mengukur seberapa besar rasio antara keragaman data yang mampu dijelaskan oleh model dibandingkan dengan total keragaman data, dengan persamaan sebagai berikut:

$$R_{y_{\ell}}^2 = 1 - \frac{\sum_{i=1}^n (y_{\ell i} - \hat{f}_{\ell i})^2}{\sum_{i=1}^n (y_{\ell i} - \bar{y}_{\ell})^2}, \ell = 1, 2. \quad (2.35)$$

Model yang baik adalah model yang memiliki nilai $RMSE_{y_1}$ dan $RMSE_{y_2}$ yang kecil serta nilai $R_{y_1}^2$ dan $R_{y_2}^2$ yang besar.

Hal yang perlu diperhatikan dalam pemodelan data adalah bagaimana membentuk model yang sesuai/tepat menggambarkan fenomena riil yang kompleks bukan mendapatkan model yang sempurna. Salah satu cara untuk mengetahuinya adalah dengan melakukan validasi model untuk melihat ketepatan model dalam prediksi. Terdapat beberapa metode validasi, diantaranya membagi data menjadi dua bagian yaitu data *training* dan data *testing*. Data training digunakan untuk membangun model, sedangkan data testing digunakan untuk menguji model (Jekabsons, 2015 dan Milborrow, 2015). Salah satu ukuran yang dapat digunakan untuk validasi model adalah dengan menentukan nilai *RMSEP*. Rumus *RMSEP* dinyatakan sebagai berikut:

$$RMSEP_{y_{\ell}} = \sqrt{\frac{\sum_{i=1}^{n_{testing}} (y_{\ell i} - \hat{f}_{\ell i})^2}{n_{testing}}}, \ell = 1, 2. \quad (2.36)$$

Nilai *RMSEP* dapat dijadikan indikasi kelayakan model. Semakin kecil nilai *RMSEP* maka semakin baik prediksi model yang dihasilkan.

2.6. Uji Linieritas Data dengan *Reset Test*

Pada penelitian ini menekankan pemodelan regresi nonparametrik yang beranjak dari bentuk model yang tidak memenuhi asumsi linieritas atau dengan kata lain antara prediktor dengan respon memiliki bentuk hubungan yang tidak linier (Wu dan Zhang, 2006). Salah satu alat uji untuk mengetahui hubungan linieritas adalah *Regression Specification Error Test* atau disingkat *RESET test* yang pertama kali diperkenalkan oleh Ramsey pada tahun 1969. Langkah-langkah dalam *RESET test* adalah sebagai berikut:

(i). Misalkan diberikan model persamaan regresi sebagai berikut:

$$y_i = \alpha_0 + \alpha_1 x_{1i} + \alpha_2 x_{2i} + \dots + \alpha_p x_{pi} + \varepsilon_i. \quad (2.37)$$

Cari estimasi parameter model dalam persamaan (2.37) dengan metode OLS, kemudian hitung:

$$\hat{y}_i = \hat{\alpha}_0 + \hat{\alpha}_1 x_{1i} + \hat{\alpha}_2 x_{2i} + \dots + \hat{\alpha}_p x_{pi}, \quad (2.38)$$

dan tentukan nilai R^2 yang disebut dengan R_{lama}^2 sebagai berikut:

$$R_{lama}^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}. \quad (2.39)$$

(ii). Tambahkan \hat{y}_i yang diperoleh dari persamaan (2.38) kedalam persamaan (2.37), sehingga diperoleh model persamaan regresi baru sebagai berikut:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + \beta_{p+1} \hat{y}_i^2 + \beta_{p+2} \hat{y}_i^3 + \varepsilon_i. \quad (2.40)$$

Cari estimasi parameter model baru dalam persamaan (2.40) dengan metode OLS, kemudian hitung:

$$\hat{y}_i^* = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \dots + \hat{\beta}_p x_{pi} + \hat{\beta}_{p+1} \hat{y}_i^2 + \hat{\beta}_{p+2} \hat{y}_i^3, \quad (2.41)$$

dan tentukan nilai R^2 yang disebut dengan R_{baru}^2 sebagai berikut:

$$R_{baru}^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i^*)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}. \quad (2.42)$$

(iii). Lakukan uji F untuk mengetahui apakah variabel tambahan pada persamaan (2.40) memiliki pengaruh yang signifikan dengan hipotesis sebagai berikut:

$$H_0: \beta_{p+1} = \beta_{p+2} = 0$$

$$H_1: \text{Minimal ada satu } \beta_j \neq 0, \text{ untuk } j = p+1, p+2.$$

Statistik uji yang digunakan adalah sebagai berikut:

$$F = \frac{(R_{baru}^2 - R_{lama}^2) / df1}{(1 - R_{baru}^2) / df2}, \text{ dimana:}$$

$df1$ = jumlah variabel bebas yang ditambahkan pada model baru,

$df2 = n$ - jumlah parameter pada model baru.

Jika $F > F_{(1-\alpha; df1; df2)}$ atau $p\text{-value} < \alpha$, dengan α adalah taraf signifikan yang dipilih, maka H_0 ditolak, atau dapat disimpulkan bahwa model pada persamaan (2.37) adalah tidak linier.

2.7. Uji Normal Multivariat Data dengan Q-Q Plot

Vektor $\tilde{x}' = (x_1, x_2, \dots, x_p)$ dikatakan berdistribusi normal multivariat dengan *mean* μ dan *covarians* Σ jika mempunyai *probability density function* :

$$f(x_1, x_2, \dots, x_p) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \exp \left[-\frac{1}{2} (\tilde{x} - \mu)' \Sigma (\tilde{x} - \mu) \right] \quad (2.43)$$

Jika $\tilde{x}' = (x_1, x_2, \dots, x_p)$ berdistribusi normal multivariat maka

$(\tilde{x} - \mu)' \Sigma^{-1} (\tilde{x} - \mu)$ berdistribusi χ_p^2 . Berdasarkan sifat ini maka pemeriksaan distribusi normal multivariat dapat dilakukan dengan cara membuat q - q plot dari nilai $d_i^2 = (\tilde{x}_i - \bar{\tilde{x}})' \mathbf{S}^{-1} (\tilde{x}_i - \bar{\tilde{x}})$, $i = 1, 2, \dots, n$.

Tahapan dari pembuatan q - q plot ini adalah sebagai berikut (Johnson, 2007) :

1. Tentukan nilai vektor rata-rata $\bar{\tilde{x}}$.
2. Tentukan nilai matriks kovarians \mathbf{S}^{-1} .
3. Tentukan nilai jarak Mahalanobis setiap titik observasi dengan vektor rata-ratanya: $d_i^2 = (\tilde{x}_i - \bar{\tilde{x}})' \mathbf{S}^{-1} (\tilde{x}_i - \bar{\tilde{x}})$, $i = 1, 2, \dots, n$.

4. Urutkan nilai d_i^2 dari nilai terkecil ke nilai terbesar : $d_{(1)}^2 \leq d_{(2)}^2 \leq \dots \leq d_{(n)}^2$.
5. Tentukan nilai $k_i = \frac{i-0,5}{n}$, $i = 1, 2, \dots, n$.
6. Untuk distribusi normal standar, tentukan nilai q_i sedemikian sehingga
$$\int_{-\infty}^{q_i} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right) dz = k_i.$$
7. Buat *scatter plot* antara $d_{(i)}^2$ dengan q_i .
8. Tentukan nilai t , dimana t adalah persentase banyaknya nilai $d_i^2 \leq \chi_{p;0,50}^2$.
9. Jika *scatter plot* cenderung membentuk garis lurus dan nilai $t > 50\%$, maka data dinyatakan berdistribusi normal multivariat.

2.8. Teorema dan Proposisi Matriks Aljabar

Sub bab berikut ini menyajikan beberapa teorema dan proposisi tentang matriks aljabar. Teorema dan proposisi tersebut akan dipergunakan pada pembahasan teori (Bab 4) untuk mendapatkan estimasi fungsi regresi dari model MARS birespon kontinu yang akan dikembangkan. Bukti teorema dan proposisi tersebut dapat dibaca pada pustaka yang diacu.

Teorema 2.8.1 (Gupta dan Nagar, 2000)

Jika \mathbf{A} dan \mathbf{B} masing-masing adalah matriks nonsingular berorde $m \times m$ dan $n \times n$, maka:

- (i). $|\mathbf{A} \otimes \mathbf{B}| = |\mathbf{A}|^n |\mathbf{B}|^m$,
- (ii). $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$.

Proposisi 2.8.1 (Searle, Casella dan McCulloch, 2006)

Jika α adalah vektor berorde $p \times 1$ dan \mathbf{A} adalah matriks berorde $p \times m$ maka:

$$\frac{\partial}{\partial \alpha} (\alpha' \mathbf{A}) = \mathbf{A},$$

Proposisi 2.8.2 (Searle *et al.*, 2006)

Jika α adalah vektor berorde $p \times 1$ dan \mathbf{A} adalah matriks simetris berorde $p \times p$ maka:

$$\frac{\partial}{\partial \underline{\alpha}} (\underline{\alpha}' \mathbf{A} \underline{\alpha}) = 2 \mathbf{A} \underline{\alpha},$$

Proposisi 2.8.3 (Broxson, 2006)

Jika $\mathbf{A}, \mathbf{B}, \mathbf{C}$ dan \mathbf{D} masing-masing adalah matriks berorde $p \times m$, $m \times n$, $p \times q$ dan $q \times n$, maka:

$$(\text{vec}(\mathbf{A}))' (\mathbf{B} \otimes \mathbf{C}) \text{vec}(\mathbf{D}) = \text{tr}(\mathbf{B}' \mathbf{A}' \mathbf{C} \mathbf{D}).$$

Proposisi 2.8.4 (Christensen, 1991)

Jika $\mathbf{A}(t)$ adalah matriks berorde $r \times r$, dengan elemen dari $\mathbf{A}(t)$ adalah fungsi dari skalar t , maka:

$$\frac{\partial}{\partial t} \text{tr}[\mathbf{A}(t)] = \text{tr} \left[\frac{\partial \mathbf{A}(t)}{\partial t} \right].$$

Proposisi 2.8.5 (Searle *et al.*, 2006)

Jika $\mathbf{A}(t)$ adalah matriks nonsingular, dengan elemen dari $\mathbf{A}(t)$ adalah fungsi dari skalar t , maka:

- (i). $\frac{\partial \mathbf{A}(t)^{-1}}{\partial t} = -\mathbf{A}(t)^{-1} \frac{\partial \mathbf{A}(t)}{\partial t} \mathbf{A}(t)^{-1},$
- (ii). $\frac{\partial}{\partial t} \ln |\mathbf{A}(t)| = \text{tr} \left(\mathbf{A}(t)^{-1} \frac{\partial \mathbf{A}(t)}{\partial t} \right).$

2.9. Definisi Operasional Variabel Penelitian

Tujuan akhir pembangunan adalah kesejahteraan rakyat. Manusia bukan hanya merupakan obyek pembangunan tetapi diharapkan dapat menjadi subyek, sehingga dapat memberikan kontribusi yang bermanfaat bagi kemajuan suatu wilayah yang secara makro menjadi kemajuan suatu negara. Keberhasilan pembangunan dapat diukur dengan beberapa indikator diantaranya adalah penurunan persentase penduduk miskin dan peningkatan IPM. Naik turunnya nilai dari kedua indikator tersebut dapat diteliti dengan beberapa prediktor yang diduga dapat mempengaruhinya diantaranya adalah angka pertumbuhan ekonomi, angka beban ketergantungan, persentase pengeluaran perkapita untuk makanan, tingkat

pengangguran terbuka (TPT) dan angka persentase penduduk yang pekerjaannya utamanya di sektor pertanian.

2.9.1. Persentase Penduduk Miskin (*Head Count Index*)

Kemiskinan secara etimologis berasal dari kata “miskin” yang artinya tidak berharta benda dan serba kekurangan. Terdapat banyak peneliti yang mengkaji tentang teori-teori kemiskinan dan bentuk-bentuk kemiskinan. Beberapa diantaranya adalah Chambers (1998), Sumarjan (1993) dan Tambunan (2001).

- Chambers (1983), menyatakan bahwa kemiskinan adalah suatu *integrated concept* yang memiliki lima dimensi, yaitu : (1) kemiskinan itu sendiri, (2) kelemahan fisik, (3) keterasingan, (4) kerentanan dan (5) ketidakberdayaan.
- Sumarjan (1993), membagi kemiskinan berdasarkan penyebabnya, yaitu kemiskinan kultural dan kemiskinan struktural.
 - a. Kemiskinan kultural, adalah kemiskinan yang umumnya diakibatkan oleh mental atau nilai-nilai yang dianut seseorang, seperti malas bekerja, mudah menyerah pada nasib dan kurangnya kerja keras untuk memperbaiki tingkat kesejahteraannya, meskipun ada usaha dari pihak lain untuk membantunya.
 - b. Kemiskinan struktural, adalah kemiskinan yang muncul bukan karena ketidakmampuan atau kemalasan seseorang untuk bekerja, melainkan karena ketidakmampuan sistem dan struktur sosial dalam menyediakan kesempatan-kesempatan yang memungkinkan orang miskin dapat bekerja.
- Tambunan (2001), menyatakan bahwa umumnya terdapat dua indikator untuk mengukur tingkat kemiskinan, yaitu kemiskinan absolut dan kemiskinan relatif.
 - a. Kemiskinan absolut, mengukur kemiskinan berdasarkan ketidakmampuan seseorang dengan pendapatan yang diperolehnya untuk mencukupi kebutuhan dasar minimum yang diperlukan untuk hidup setiap hari.
 - b. Kemiskinan relatif, mengukur kemiskinan berdasarkan ketidakmampuan untuk mencapai standar kehidupan yang ditetapkan masyarakat setempat, dimana penentuannya sangat subyektif.

Di Indonesia, Badan Pusat Statistik (BPS) mengukur kemiskinan berdasarkan teori kemiskinan absolut, yang diukur berdasarkan ketidakmampuan seseorang untuk mencukupi kebutuhan pokok minimum. Salah satu indikator

yang digunakan adalah persentase penduduk miskin atau *Head Count Index* (HCI) yang menunjukkan persentase jumlah penduduk yang hidup dibawah garis kemiskinan terhadap total jumlah penduduk di suatu wilayah. Indikator ini dapat melihat kecenderungan tingkat kemiskinan di suatu wilayah sehingga mempermudah perbandingan antar wilayah maupun negara. BPS menghitung persentase penduduk miskin menggunakan garis kemiskinan yaitu suatu ukuran yang menyatakan besarnya pengeluaran untuk memenuhi kebutuhan dasar minimum makanan dan kebutuhan non makanan, atau standar yang menyatakan batas seseorang dikatakan miskin bila dipandang dari sudut konsumsi (BPS, 2013). Foster, Greer, dan Thorbecke (1984) memberikan formula persentase penduduk miskin sebagai berikut :

$$HCI_i = \frac{q_i}{n_i}.$$

Keterangan :

HCI_i : persentase penduduk miskin di wilayah ke- i ,

q_i : banyaknya penduduk hidup di bawah garis kemiskinan pada wilayah ke- i ,

n_i : banyaknya penduduk pada wilayah ke- i .

2.9.2. Indeks Pembangunan Manusia

Keberhasilan pembangunan diukur dengan beberapa parameter, diantaranya adalah Indeks Pembangunan Manusia (IPM) atau *Human Development Index* (HDI). Alat ukur ini dikembangkan oleh Mahbub ul Haq dalam bukunya yang berjudul *Reflections on Human Development* dan telah disepakati dunia melalui *United Nation Development Program* (UNDP) pada tahun 1990. Sebagai alat indikator, IPM bermanfaat untuk:

1. Mengukur keberhasilan dalam upaya membangun kualitas hidup manusia (masyarakat/penduduk).
2. Menentukan peringkat atau level pembangunan di suatu wilayah / daerah.

Sejak tahun 2010, UNDP menyempurnakan metode IPM sebelumnya dengan merubah salah satu indikator penyusun IPM yaitu angka melek huruf menjadi angka harapan lama sekolah, dengan alasan angka melek huruf di sebagian besar daerah sudah tinggi, sehingga tidak dapat membedakan tingkat

pendidikan antar daerah dengan baik. Selain itu UNDP, juga menyempurnakan metode penghitungan IPM dengan merubah penggunaan rumus rata-rata aritmatik dengan rata-rata geometrik, dengan alasan bahwa capaian satu dimensi penyusun IPM seharusnya tidak ditutupi oleh capaian dimensi lain. Artinya, untuk mewujudkan pembangunan manusia yang baik, semua dimensi penyusun IPM (kesehatan, pendidikan dan pengeluaran) harus memperoleh perhatian yang sama besar karena sama pentingnya.

Saat ini penghitungan IPM di Indonesia, dilakukan oleh BPS dengan menyesuaikan metode penghitungan baru yang ditetapkan oleh UNDP. BPS menghitung IPM sebagai rata-rata geometrik dari indeks kesehatan, indeks pendidikan dan indeks standar hidup layak. Indeks kesehatan dihitung berdasarkan jumlah anak lahir hidup (*children ever born: CEB*) dan jumlah anak masih hidup (*children surviving: CS*) dari wanita usia 15 – 49 tahun. Indeks pendidikan diukur berdasarkan angka harapan lama sekolah dari penduduk usia 7 tahun keatas dan angka rata-rata lama sekolah dari penduduk usia 25 tahun keatas. Sedangkan indeks standar hidup layak diukur dengan rata-rata pengeluaran konsumsi riil per kapita yang disesuaikan pertahun. Setiap komponen IPM distandarisasi dengan nilai minimum dan maksimum sebelum digunakan untuk menghitung IPM. Setelah distandarisasi, angka IPM dihitung dengan formula:

$$IPM_i = \sqrt[3]{IS_i \times IP_i \times IL_i} \times 100,$$

dengan :

$$IP_i = \frac{IHLS_i + IRLS_i}{2}.$$

Keterangan :

IPM_i : indeks pembangunan manusia pada wilayah ke- i ,

IS_i : indeks kesehatan pada wilayah ke- i ,

IP_i : indeks pendidikan pada wilayah ke- i ,

IL_i : indeks standar hidup layak pada wilayah ke- i ,

$IHLS_i$: indeks harapan lama sekolah pada wilayah ke- i ,

$IRLS_i$: indeks rata-rata lama sekolah pada wilayah ke- i .

Angka IPM berkisar antara 0 hingga 100, semakin mendekati nilai 100 merupakan indikasi pembangunan manusia yang semakin baik. UNDP membagi status pembangunan manusia suatu wilayah atau negara ke dalam empat golongan berdasarkan nilai IPM, yaitu: rendah ($IPM < 60$), sedang/menengah ($60 \leq IPM < 70$), tinggi ($70 \leq IPM < 80$) dan sangat tinggi ($IPM \geq 80$).

Terdapat beberapa penelitian yang mengkaji hubungan timbal balik antara IPM dengan persentase penduduk miskin, diantaranya adalah Madan (2012) dan Septiana (2012), dengan penelitian masing-masing di India dan di Indonesia. Penelitian tersebut menyatakan bahwa terdapat korelasi negatif antara IPM dengan kemiskinan, dimana semakin tinggi IPM di negara tersebut maka tingkat kemiskinannya akan semakin rendah. Hal ini disebabkan karena dengan meningkatnya kesejahteraan masyarakat (tidak miskin) maka kemampuan untuk memenuhi kebutuhan hidup seperti pendidikan, kesehatan dan kehidupan yang layak (komponen penyusun IPM) menjadi semakin baik, sehingga berdampak terhadap peningkatan angka IPM. Sebaliknya dengan semakin sehat dan baiknya tingkat pendidikan masyarakat (komponen IPM), maka produktivitas kerja akan meningkat dan pendapatan ikut meningkat, sehingga tidak miskin.

2.9.3. Pertumbuhan Ekonomi

Pertumbuhan ekonomi merupakan suatu proses perubahan kondisi perekonomian suatu wilayah secara berkesinambungan menuju keadaan yang lebih baik dari tahun ke tahun. Indikator ini biasanya digunakan untuk melihat sampai seberapa jauh keberhasilan pembangunan di suatu wilayah dalam jangka waktu tertentu. Hal ini berarti bahwa indikator ini dapat dipakai pula untuk menentukan arah kebijaksanaan pembangunan yang akan datang. Menurut Todaro dan Smith (2006), ada tiga faktor utama dalam pertumbuhan ekonomi yaitu: (1) Akumulasi modal termasuk semua investasi baru yang berwujud tanah, peralatan fiskal dan sumber daya manusia, (2) Pertumbuhan penduduk dan angkatan kerja dan (3) Kemajuan teknologi.

BPS menghitung pertumbuhan ekonomi di Indonesia dari data Produk Domestik Regional Bruto (PDRB) atas dasar harga konstan yaitu nilai tambah barang dan jasa yang dihasilkan dari seluruh kegiatan perekonomian di suatu

wilayah yang dihitung menggunakan harga pada tahun tertentu sebagai tahun dasar, dimana saat ini menggunakan tahun dasar 2010. Pertumbuhan yang positif menunjukkan adanya peningkatan perekonomian dan sebaliknya. BPS (2013) menyatakan bahwa pertumbuhan ekonomi dapat dihitung dengan formula:

$$PE_{t,i} = \frac{PDRB_{t,i} - PDRB_{t-1,i}}{PDRB_{t-1,i}} \times 100\%.$$

Keterangan :

$PE_{t,i}$: Pertumbuhan ekonomi pada tahun t di wilayah ke- i ,

$PDRB_{t,i}$: PDRB pada tahun t di wilayah ke- i ,

$PDRB_{t-1,i}$: PDRB pada tahun $t-1$ (sebelumnya) di wilayah ke- i .

Ranis *et al.* (2000) dalam penelitiannya mengkaji bahwa pertumbuhan ekonomi memberi pengaruh terhadap angka IPM beberapa negara berkembang di dunia, dimana semakin tinggi pertumbuhan ekonomi di suatu negara akan berpengaruh terhadap tingginya angka IPM di negara tersebut. Hal ini disebabkan karena dengan meningkatnya pertumbuhan ekonomi maka pembiayaan akan pelayanan kesehatan dan pendidikan oleh pemerintah menjadi semakin baik, sehingga angka IPM meningkat juga. Selain itu pertumbuhan ekonomi akan mendorong peningkatan pendapatan perkapita masyarakat sehingga angka IPM juga mengalami peningkatan. Son *et al.* (2004) juga mengkaji bahwa pertumbuhan ekonomi memberikan pengaruh terhadap angka kemiskinan dengan studi kasus pada beberapa negara di Asia. Penelitian tersebut menyatakan bahwa penurunan kemiskinan dapat dicapai dengan meningkatnya pertumbuhan ekonomi. Peneliti lainnya yaitu Ijaiya *et al.* (2011) yang mengkaji pengaruh pertumbuhan ekonomi terhadap penurunan kemiskinan di Nigeria, yang menyatakan bahwa pertumbuhan ekonomi dapat membuka lapangan pekerjaan dan kegiatan produktif bagi penduduk miskin.

2.9.4. Angka Beban Ketergantungan

Angka beban ketergantungan dapat digunakan sebagai indikator yang secara kasar dapat menunjukkan keadaan ekonomi suatu wilayah apakah tergolong wilayah maju atau wilayah yang sedang berkembang. Angka beban ketergantungan merupakan salah satu indikator demografi yang penting. Semakin

tingginya persentase angka beban ketergantungan menunjukkan semakin tingginya beban yang harus ditanggung penduduk yang produktif untuk membiayai hidup penduduk yang belum produktif dan tidak produktif lagi. Angka beban ketergantungan diperoleh dengan membandingkan antara banyaknya penduduk pada kelompok usia produktif (15-64 tahun) terhadap kelompok usia muda (kurang dari 15 tahun) dan usia tua (65 tahun keatas). Penduduk muda berusia dibawah 15 tahun umumnya dianggap sebagai penduduk yang belum produktif karena secara ekonomis masih tergantung pada orang tua atau orang lain yang menanggungnya. Selain itu, penduduk berusia diatas 65 tahun juga dianggap tidak produktif lagi sesudah melewati masa pensiun, pada masa ini orang tersebut kembali menjadi beban tanggungan hidup bagi orang lain. Angka beban ketergantungan diperoleh dengan formula sebagai berikut:

$$ABT_i = \frac{P_i(<15) + P_i(65+)}{P_i(15-64)} \times 100\% .$$

Keterangan:

ABT_i : Angka beban tanggungan pada wilayah ke- i ,

$P_i(<15)$: Jumlah penduduk berusia kurang dari 15 tahun pada wilayah ke- i ,

$P_i(15-64)$: Jumlah penduduk usia 15 sampai dengan 64 tahun pada wilayah ke- i ,

$P_i(65+)$: Jumlah penduduk berusia 65 tahun keatas pada wilayah ke- i .

Terdapat beberapa penelitian yang telah mengkaji pengaruh angka beban ketergantungan terhadap angka persentase penduduk miskin dan IPM. Peneliti tersebut diantaranya adalah Chaudhuri (2003) yang mengkaji pengaruh angka beban ketergantungan penduduk terhadap kemiskinan di Kolumbia dan Pratowo (2011) yang mengkaji pengaruh angka beban ketergantungan terhadap IPM di Propinsi Jawa Tengah. Hasil penelitian menyatakan bahwa semakin besar angka beban ketergantungan, maka semakin rendah angka IPM dan semakin tinggi persentase penduduk miskin. Hal ini dapat dipahami karena semakin besar angka beban tanggungan, maka semakin berat beban tanggungan penduduk usia produktif untuk menanggung kebutuhan (pendidikan, kesehatan, kehidupan layak dan lain-lain) penduduk usia tidak produktif, sehingga dampaknya IPM menurun dan kemiskinan meningkat.

2.9.5. Persentase Pengeluaran Perkapita untuk Makanan

Proporsi pengeluaran perkapita untuk konsumsi makanan terhadap seluruh pengeluaran rumah tangga merupakan salah satu indikator kesejahteraan di suatu wilayah. Pengeluaran rumah tangga dapat digolongkan kedalam pengeluaran untuk konsumsi makanan dan pengeluaran untuk konsumsi non makanan. Rumah tangga dengan proporsi pengeluaran yang lebih besar untuk konsumsi non makanan mengindikasikan rumah tangga yang berpenghasilan tinggi. Rumah tangga dengan kondisi sebaliknya yaitu dengan pendapatan terbatas biasanya memiliki proporsi pengeluaran lebih kecil untuk konsumsi non makanan. Hal ini disebabkan karena sebagian besar penghasilannya didahulukan untuk pemenuhan konsumsi makanan dan jika ada sisanya baru digunakan untuk konsumsi non makanan. Dengan kata lain dapat dikatakan bahwa rumah tangga akan semakin sejahtera bila persentase pengeluaran untuk konsumsi makanan jauh lebih kecil dibandingkan persentase pengeluaran untuk konsumsi non makanan. Hal ini berarti bahwa tinggi rendahnya tingkat pendapatan masyarakat di suatu wilayah dapat diprediksi dengan besarnya proporsi pengeluaran perkapita untuk konsumsi makanan (BPS, 2013), yang dihitung dengan formula sebagai berikut:

$$Q_i = \frac{PM_i}{PT_i}.$$

Keterangan :

Q_i : persentase pengeluaran perkapita per bulan untuk makanan di wilayah ke- i (%),

PM_i : pengeluaran perkapita per bulan untuk makanan di wilayah ke- i (Rp/bulan),

PT_i : pengeluaran perkapita per bulan untuk makanan dan non makanan (total) di wilayah ke- i (Rp/bulan).

Terdapat beberapa penelitian yang telah mengkaji pengaruh persentase pengeluaran perkapita makanan terhadap angka IPM dan persentase penduduk miskin di Indonesia. Penelitian tersebut diantaranya adalah Pratowo (2011) yang mengkaji pengaruh pola konsumsi non makanan terhadap IPM. Wulandari dan Budiantara (2014) yang mengkaji pengaruh pengeluaran perkapita makanan pada

persentase penduduk miskin di Jawa Timur dengan pendekatan spline. Hasil penelitian menyatakan bahwa semakin besar angka persentase pengeluaran perkapita makanan maka semakin rendah angka IPM dan semakin tinggi persentase penduduk miskin. Hal ini dapat dipahami karena semakin besarnya porsi pengeluaran makanan dalam kebutuhan masyarakat, maka porsi pembiayaan untuk memenuhi kebutuhan non makanan (pendidikan, kesehatan, kepemilikan barang dan lain-lain) semakin kurang, dampaknya angka IPM turun dan kemiskinan meningkat.

2.9.6. Tingkat Pengangguran Terbuka (TPT)

Pengangguran merupakan salah satu masalah yang selalu dihadapi dan sulit dihindari oleh suatu negara, baik di negara berkembang maupun negara maju. Todaro (2003), menyatakan bahwa pengangguran dapat dibedakan menjadi tiga:

1. Pengangguran Terselubung (*Disguised Unemployment*), adalah tenaga kerja yang tidak bekerja secara optimal karena suatu alasan tertentu.
2. Setengah penganggur (*Under Employment*) adalah tenaga kerja yang tidak bekerja secara optimal karena tidak ada lapangan pekerjaan, biasanya tenaga kerja setengah menganggur ini merupakan tenaga kerja yang bekerja kurang dari 35 jam selama seminggu.
3. Pengangguran terbuka (*Open Unemployment*) adalah tenaga kerja yang sungguh-sungguh tidak memiliki pekerjaan. Pengangguran jenis ini cukup banyak karena memang belum mendapat pekerjaan padahal telah berusaha maksimal.

Indikator pengangguran yang digunakan oleh BPS adalah tingkat pengangguran terbuka (TPT). BPS menghitung angka TPT dengan membandingkan jumlah pengangguran dengan jumlah angkatan kerja (penduduk yang bekerja dan penduduk yang menganggur) di suatu wilayah. Pengangguran yang dihitung dalam TPT adalah penduduk usia kerja (15 tahun keatas) yang sedang mencari kerja atau sedang mempersiapkan usaha, atau tidak mencari kerja karena merasa tidak mungkin memperoleh pekerjaan, atau sudah diterima kerja tetapi belum mulai kerja (BPS, 2013). Sementara konsep bekerja yang dimaksud adalah kegiatan melakukan pekerjaan dengan maksud memperoleh atau

membantu memperoleh penghasilan atau keuntungan paling sedikit selama satu jam dalam seminggu yang lalu. Bekerja selama satu jam tersebut harus dilakukan berturut-turut dan tidak terputus. BPS (2013) menyatakan bahwa angka TPT diperoleh dengan formula sebagai berikut:

$$TPT_i = \frac{P_i}{AK_i} \times 100\% .$$

Keterangan:

TPT_i : tingkat pengangguran terbuka di wilayah ke- i ,

P_i : banyaknya pengangguran di wilayah ke- i ,

AK_i : banyaknya angkatan kerja di wilayah ke- i .

Tambunan (2001), menyatakan bahwa pengangguran dapat mempengaruhi tingkat kemiskinan dengan berbagai cara, antara lain:

1. Jika rumah tangga memiliki batasan likuiditas, yang berarti bahwa konsumsi saat ini sangat dipengaruhi oleh pendapatan saat ini, maka bencana pengangguran akan secara langsung mempengaruhi *income poverty rate* dengan *consumption poverty rate*.
2. Jika rumah tangga tidak menghadapi batasan likuiditas, yang berarti bahwa konsumsi saat ini tidak terlalu dipengaruhi oleh pendapatan saat ini, maka peningkatan pengangguran akan menyebabkan peningkatan kemiskinan dalam jangka panjang, tetapi tidak terlalu berpengaruh dalam jangka pendek.

Selain itu, Sukirno (2004) menyatakan bahwa efek buruk dari pengangguran adalah mengurangi pendapatan masyarakat yang pada akhirnya mengurangi tingkat kemakmuran yang dicapai seseorang. Semakin turunnya kesejahteraan masyarakat karena menganggur tentunya akan meningkatkan peluang mereka terjebak dalam kemiskinan karena tidak memiliki pendapatan sehingga berdampak terhadap turunnya kualitas hidup mereka, sehingga angka IPM juga ikut turun.

2.9.7. Persentase Penduduk yang Pekerjaan Utamanya di Sektor Pertanian

Data tentang distribusi sektoral penyerapan tenaga kerja dapat digunakan sebagai salah satu indikator untuk melihat kemampuan sektor-sektor ekonomi

dalam menyerap tenaga kerja dan juga sebagai tolak ukur kemajuan perekonomian suatu daerah. Tahapan kemajuan perekonomian suatu daerah dari tradisional menuju daerah industri, salah satunya ditandai dengan adanya transformasi sektoral tenaga kerja dari sektor primer (pertanian) dengan produktivitas rendah ke sektor-sektor dengan produktivitas lebih tinggi yaitu sektor sekunder (industri, konstruksi, listrik, gas dan air) dan tersier (perdagangan, transportasi, komunikasi, keuangan dan jasa). Sehingga, persentase tenaga kerja di sektor primer (pertanian) akan menurun dan sebaliknya pada sektor sekunder dan tersier akan meningkat. Banyaknya persentase penduduk yang bekerja di sektor pertanian dapat dihitung dengan formula sebagai berikut:

$$T_i = \frac{JT_i}{JB_i} \times 100\%.$$

Keterangan:

T_i : persentase penduduk yang bekerja di sektor pertanian pada wilayah ke- i ,

JT_i : jumlah penduduk yang bekerja di sektor pertanian pada wilayah ke- i ,

JB_i : jumlah seluruh penduduk yang bekerja pada wilayah ke- i .

Terdapat penelitian yang mengkaji pengaruh penyerapan sektor tenaga kerja terhadap tingkat kemiskinan dan IPM, seperti penelitian Ayomi (2014) di wilayah Madiun dan sekitarnya tahun 2003-2012. Banyaknya persentase penduduk pertanian di suatu daerah, dapat berpengaruh terhadap tingginya persentase penduduk miskin dan rendahnya angka IPM di daerah tersebut. Hal ini dapat dipahami karena pendapatan penduduk di sektor pertanian umumnya sangat kecil, apalagi menurut hasil sensus pertanian 2013 menyatakan bahwa sebagian besar petani yang ada di Indonesia adalah petani gurem yang mengusahakan lahan pertanian kurang dari 0,5 Ha. Akibatnya produktivitas hasil pertanian mereka tidak optimal, dan tidak cukup untuk membiayai kebutuhan hidup mereka setiap hari. Dampaknya penduduk tetap miskin dan IPM rendah, karena tidak memiliki cukup biaya untuk bersekolah, berobat atau memiliki kehidupan yang layak.

BAB 3

METODOLOGI PENELITIAN

Ada dua kajian utama yang akan dilakukan dalam penelitian ini, yaitu kajian teoritik dan kajian terapan. Kajian teoritik difokuskan pada pengembangan metode MARS dalam regresi nonparametrik birespon kontinu. Kajian ini meliputi estimasi fungsi regresi dan estimasi bobot yang mengakomodasi adanya korelasi antar respon dalam model MARS birespon kontinu. Kajian terapan difokuskan pada penerapan hasil kajian teoritik pada data simulasi dan pada data riil. Penyelesaian dari kajian-kajian tersebut dilakukan berdasarkan langkah-langkah penelitian yang diuraikan dalam bab ini.

3.1. Estimasi Fungsi Regresi dalam Model MARS Birespon Kontinu

Sub-bab ini merupakan bagian dari langkah kajian teoritik yang bertujuan untuk mendapatkan estimasi fungsi regresi dalam model MARS birespon kontinu. Langkah-langkah untuk mendapatkan estimasi fungsi regresi ini sebagai berikut:

1. Menetapkan bentuk umum model regresi untuk dua respon kontinu dan p prediktor pada n pengamatan yang mengikuti model regresi nonparametrik birespon kontinu sebagai berikut:

$$y_{\ell i} = f_{\ell}(x_{1i}, x_{2i}, \dots, x_{pi}) + \varepsilon_{\ell i}, \quad \ell = 1, 2; \quad i = 1, 2, \dots, n,$$

dengan *random error* $\varepsilon_{\ell i}$ diasumsikan berdistribusi normal, dengan:

$$E(\varepsilon_{\ell i}) = 0, \quad E(\varepsilon_{\ell i} \varepsilon_{\ell i'}) = \begin{cases} \sigma_{\ell\ell}, & \text{jika } i = i' \\ 0, & \text{jika } i \neq i' \end{cases},$$

$$\text{dan } E(\varepsilon_{1i} \varepsilon_{2i'}) = \begin{cases} \sigma_{12}, & \text{jika } i = i' \\ 0, & \text{jika } i \neq i' \end{cases}.$$

2. Menghampiri kurva regresi f_{ℓ} pada langkah 1 dengan fungsi MARS yang dinyatakan dalam persamaan sebagai berikut:

$$f_{\ell}(x_{1i}, x_{2i}, \dots, x_{pi}) = \alpha_{\ell 0} + \sum_{m_{\ell}=1}^{M_{\ell}} \alpha_{\ell m_{\ell}} \prod_{k_{\ell}=1}^{K_{m_{\ell}}} \left[S_{k_{\ell} m_{\ell}}(x_{v(k_{\ell}, m_{\ell})i} - t_{k_{\ell} m_{\ell}}) \right]_+; \quad \ell = 1, 2; \quad i = 1, 2, \dots, n,$$

$$\text{dengan } x_{v(k_{\ell}, m_{\ell})} \in \{x_j\}_{j=1}^p \text{ dan } t_{k_{\ell} m_{\ell}} \in \{x_{v(k_{\ell}, m_{\ell})i}\}_{i=1}^n.$$

3. Menyusun persamaan regresi yang terbentuk dari langkah 1 dan 2 dalam bentuk matriks, sehingga diperoleh model MARS birespon kontinu:

$$\tilde{y} = \tilde{f}(\tilde{x}) + \varepsilon; \text{ dimana } \tilde{f}(\tilde{x}) = \mathbf{B}(\tilde{x}, \tilde{t})\alpha,$$

dengan *random error* ε diasumsikan berdistribusi normal dengan *mean* 0 dan matriks kovarians *error* \mathbf{W} .

4. Mendapatkan fungsi *likelihood* $L(\alpha|\mathbf{W}, y)$ dari model MARS birespon kontinu yang diperoleh dari langkah 3. Untuk mendapatkan pendugaan parameter model, matriks \mathbf{W} diasumsikan diketahui nilainya (*fixed variable*).
5. Menyelesaikan optimasi fungsi *likelihood* untuk mendapatkan estimasi parameter α , dengan cara menurunkan fungsi *log likelihood* pada langkah 4 terhadap α dan menyamakan hasilnya dengan 0.
6. Mendapatkan estimasi fungsi regresi MARS birespon kontinu $\tilde{f}(\tilde{x})$ berdasarkan estimasi parameter α yang diperoleh pada langkah 5.

3.2. Estimasi Matriks Bobot

Estimasi fungsi regresi MARS birespon kontinu memuat matriks kovarians *error* \mathbf{W} yang berfungsi sebagai matriks bobot dalam mengakomodasi adanya korelasi antara respon. Pada umumnya matriks \mathbf{W} tidak diketahui sehingga perlu diestimasi. Langkah-langkah untuk mendapatkan estimasi matriks \mathbf{W} sebagai berikut:

1. Menyatakan matriks \mathbf{W} berdasarkan asumsi *random error* ε pada Sub-Bab

3.1 langkah 1, sehingga diperoleh $\mathbf{W} = \begin{pmatrix} \sigma_{11}\mathbf{I} & \sigma_{12}\mathbf{I} \\ \sigma_{12}\mathbf{I} & \sigma_{22}\mathbf{I} \end{pmatrix}$ atau $\mathbf{W} = \mathbf{\Sigma} \otimes \mathbf{I}$, dimana

$\mathbf{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$ adalah matriks simetris berorde 2×2 dan \mathbf{I} adalah matriks identitas berorde n .

2. Mendapatkan fungsi *likelihood* $L(\mathbf{W}|\alpha, y)$ dari model MARS birespon kontinu yang diperoleh dari Sub-Bab 3.1 langkah 3, dengan $\mathbf{W} = \mathbf{\Sigma} \otimes \mathbf{I}$ seperti diperoleh dari langkah 1.

3. Menyelesaikan optimasi fungsi *likelihood* untuk mendapatkan estimasi matriks Σ yang mengandung estimasi σ_{11}, σ_{22} dan σ_{12} , dengan cara menurunkan fungsi *log likelihood* pada langkah 2 terhadap σ_{ij} ($i, j = 1, 2$) dan menyamakan hasilnya dengan 0.
4. Mendapatkan estimasi matriks W berdasarkan estimasi matriks Σ yang diperoleh pada langkah 3.

3.3. Penerapan Model MARS Birespon Kontinu pada Data Simulasi

Sub-bab ini merupakan bagian dari langkah kajian terapan yang bertujuan untuk menerapkan hasil kajian teoritik MARS birespon kontinu pada data simulasi. Penerapan simulasi diterapkan pada contoh bentuk fungsi asli MARS dan pada berbagai bentuk fungsi percobaan, seperti fungsi trigonometri, polinomial dan eksponensial, untuk mengetahui kebaikan dari model MARS birespon kontinu yang dikembangkan. Langkah-langkah penelitian yang dilakukan dalam simulasi ini adalah sebagai berikut:

1. Menerapkan simulasi MARS birespon kontinu pada fungsi asli, dengan langkah-langkah sebagai berikut:

- a. Merancang persamaan regresi nonparametrik birespon kontinu mengikuti contoh bentuk fungsi asli MARS sebagai berikut :

$$y_{1i} = 4 + 39(x_{1i} - 0,4)_+ + 20(0,4 - x_{1i})_+ - 35(x_{1i} - 0,4)_+(x_{2i} - 0,5)_+ - 37(x_{1i} - 0,4)_+(0,5 - x_{2i})_+ + \varepsilon_{1i},$$

$$y_{2i} = 5 + 37(x_{3i} - 0,45)_+ + 18(0,45 - x_{3i})_+ - 33(0,45 - x_{3i})_+(x_{4i} - 0,7)_+ - 39(0,45 - x_{3i})_+(0,7 - x_{4i})_+ + \varepsilon_{2i}.$$

- b. Membangkitkan prediktor dari distribusi *uniform* dengan $x_{ji} \sim U(0,1)$ untuk $j = 1, 2, 3, 4$ dan $i = 1, 2, \dots, n$. Pada contoh simulasi ini, banyaknya pengamatan n dinyatakan sama dengan 125.

- c. Membangkitkan *random error* $\varepsilon = (\varepsilon'_1, \varepsilon'_2)'$ dari distribusi normal:

$$\varepsilon \sim N(0, W) \text{ dimana } W = \begin{pmatrix} \sigma_{11} \mathbf{I} & \sigma_{12} \mathbf{I} \\ \sigma_{12} \mathbf{I} & \sigma_{22} \mathbf{I} \end{pmatrix} \text{ adalah matriks simetris berorde}$$

$2n \times 2n$ dengan \mathbf{I} adalah matriks identitas berorde n . Pada contoh

simulasi ini, varians *error* σ_{11} dan σ_{22} dinyatakan sama dengan 1, sedangkan kovarians *error* σ_{12} diperoleh dari $\sigma_{12} = \rho\sqrt{\sigma_{11}\sigma_{22}}$ dengan korelasi *error* antar respon ρ dinyatakan sama dengan 0,5, serta banyaknya pengamatan n dinyatakan sama dengan 125.

- d. Mendapatkan respon $\underline{y} = (y'_1, y'_2)'$ pada langkah 1a berdasarkan bangkitan prediktor pada langkah 1b dan bangkitan *random error* pada langkah 1c.
 - e. Mengestimasi fungsi regresi $\underline{f}(\underline{x}) = (f'_1(\underline{x}), f'_2(\underline{x}))'$ dari pola hubungan antara respon dan prediktor yang masing-masing datanya diperoleh dari langkah 1d dan langkah 1b, menggunakan estimator MARS birespon kontinu yang diperoleh dari hasil kajian teoritik (Sub-Bab 3.1 dan 3.2).
 - f. Membuat visualisasi dari hasil estimasi fungsi regresi $\underline{f}(\underline{x})$ yang diperoleh pada langkah 1e.
 - g. Mengamati kebaikan MARS birespon kontinu pada berbagai kombinasi ukuran sampel dan varians *error*, dengan membandingkan hasil estimasi (parameter dan knot) dengan nilai aslinya. Hal ini dilakukan dengan cara mengulang langkah 1a sampai dengan langkah 1e dengan replikasi sebanyak 100 kali pada kombinasi ukuran sampel $n_\ell = 70, 125, 250, 500$ dan varians *error* $\sigma_{\ell\ell} = 0,1, 0,5, 1$ ($\ell = 1, 2$). Untuk melihat pengaruh besaran ukuran sampel dan varians *error* (kecil, menengah, besar) maka dalam penelitian ini ditetapkan $n_1 = n_2$ dan $\sigma_{11} = \sigma_{22}$.
 - h. Mengamati pengaruh besaran korelasi (kecil, menengah, besar) terhadap estimasi fungsi regresi MARS birespon kontinu berdasarkan nilai rasio *RMSEP* yang dihasilkan. Nilai rasio *RMSEP* diperoleh dari nilai *RMSEP* model yang tidak mengakomodir korelasi dibagi dengan nilai *RMSEP* model yang mengakomodir korelasi. Simulasi dilakukan dengan mengulang langkah 1a sampai dengan langkah 1e dengan replikasi sebanyak 100 kali pada berbagai ukuran korelasi $\rho = -0,9, -0,5, -0,3, 0, 0,3, 0,5, 0,9$.
2. Menerapkan simulasi MARS birespon kontinu pada fungsi percobaan trigonometri-polinomial sebagai berikut:

- a. Merancang persamaan regresi nonparametrik birespon kontinu dengan contoh fungsi trigonometri-polinomial sebagai berikut :

$$y_{1i} = 5 \sin(2\pi x_{1i}) - 15(0,5 - x_{2i})^2 + 6x_{1i}x_{2i} + \varepsilon_{1i},$$

$$y_{2i} = 4 \sin(2\pi x_{3i}) + 16(x_{4i} - 0,5)^2 - 7x_{3i}x_{4i} + \varepsilon_{2i}.$$

- b. Mengulang langkah 1b sampai dengan 1f pada persamaan regresi langkah 2a, untuk mengetahui kemampuan MARS birespon kontinu dalam mendapatkan hasil estimasi yang mendekati fungsi trigonometri-polinomial melalui pendekatan grafik.
- c. Mengamati pengaruh besaran korelasi (kecil, menengah, besar) berdasarkan nilai rasio *RMSEP* yang dihasilkan oleh metode MARS birespon kontinu pada fungsi percobaan trigonometri-polinomial. Simulasi dilakukan dengan replikasi sebanyak 100 kali pada berbagai ukuran korelasi $\rho = -0,9, -0,5, -0,3, 0, 0,3, 0,5, 0,9$.

3. Menerapkan simulasi MARS birespon kontinu pada fungsi percobaan eksponensial-trigonometri sebagai berikut:

- a. Merancang persamaan regresi nonparametrik birespon kontinu dengan contoh fungsi eksponensial-trigonometri sebagai berikut :

$$y_{1i} = 6 \sin(3\pi x_{1i}) + 3 \exp(-9x_{2i}) + 7x_{1i}x_{2i} + \varepsilon_{1i}$$

$$y_{2i} = 5 \sin(3\pi x_{3i}) - 2 \exp(-8x_{4i}) + 8x_{3i}x_{4i} + \varepsilon_{2i}$$

- b. Mengulang langkah 1b sampai dengan 1f pada persamaan regresi langkah 3a, untuk mengetahui kemampuan MARS birespon kontinu dalam mendapatkan hasil estimasi yang mendekati fungsi eksponensial-trigonometri melalui pendekatan grafik.
 - c. Mengamati pengaruh besaran korelasi (kecil, menengah, besar) berdasarkan nilai rasio *RMSEP* yang dihasilkan oleh metode MARS birespon kontinu pada fungsi percobaan eksponensial-trigonometri. Simulasi dilakukan dengan replikasi sebanyak 100 kali pada berbagai ukuran korelasi $\rho = -0,9, -0,5, -0,3, 0, 0,3, 0,5, 0,9$.
4. Menerapkan simulasi MARS birespon kontinu pada fungsi percobaan polinomial-eksponensial sebagai berikut:

- a. Merancang persamaan regresi nonparametrik birespon kontinu dengan contoh fungsi polinomial-eksponensial sebagai berikut :

$$y_{1i} = 15(0,5 - x_{1i})^2 + 5 \exp(-12x_{2i}) + 8x_{1i}x_{2i} + \varepsilon_{1i}$$

$$y_{2i} = 16(x_{1i} - 0,7)^2 + 4 \exp(-7x_{2i}) + 7x_{1i}x_{2i} + \varepsilon_{2i}$$

- b. Mengulang langkah 1b sampai dengan 1f pada persamaan regresi langkah 4a, untuk mengetahui kemampuan MARS birespon kontinu dalam mendapatkan hasil estimasi yang mendekati fungsi polinomial-eksponensial melalui pendekatan grafik.
- c. Mengamati pengaruh besaran korelasi (kecil, menengah, besar) berdasarkan nilai rasio *RMSEP* yang dihasilkan oleh metode MARS birespon kontinu pada fungsi percobaan polinomial-eksponensial. Simulasi dilakukan dengan replikasi sebanyak 100 kali pada berbagai ukuran korelasi $\rho = -0,9, -0,5, -0,3, 0, 0,3, 0,5, 0,9$.

3.4. Penerapan Model MARS Birespon Kontinu pada Data Riil

Sub-bab ini merupakan bagian dari langkah kajian terapan yang bertujuan untuk menerapkan hasil kajian teoritik MARS birespon kontinu pada data riil, yaitu dengan menerapkannya pada pemodelan data persentase penduduk miskin dan IPM. Sub-bab ini meliputi sumber data, variabel penelitian dan langkah-langkah analisis.

3.4.1. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder mengenai persentase penduduk miskin dan IPM, yang bersumber dari beberapa publikasi terbitan BPS tahun 2013, seperti publikasi Indikator Kesejahteraan Rakyat (INKESRA), publikasi IPM, publikasi PDRB, publikasi Data dan Informasi kemiskinan dan publikasi lainnya. Unit observasi yang digunakan dalam penelitian ini adalah seluruh kabupaten/kota yang ada di pulau Jawa.

3.4.2. Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri atas dua jenis, yaitu respon dan prediktor. Respon terdiri atas dua variabel (birespon) yang saling

berkorelasi, sedangkan prediktor terdiri atas lima variabel. Unit observasi yang digunakan untuk setiap variabel adalah wilayah kabupaten/kota.

1. Respon (y) :

y_{1i} : persentase penduduk miskin pada wilayah ke- i ,

y_{2i} : IPM pada wilayah ke- i .

2. Prediktor (x) :

x_{1i} : pertumbuhan ekonomi pada wilayah ke- i ,

x_{2i} : angka beban ketergantungan pada wilayah ke- i ,

x_{3i} : persentase pengeluaran perkapita untuk makanan pada wilayah ke- i ,

x_{4i} : tingkat pengangguran terbuka pada wilayah ke- i ,

x_{5i} : persentase penduduk yang pekerjaan utamanya di sektor pertanian pada wilayah ke- i .

3.4.3. Struktur Data Penelitian

Struktur data berdasarkan variabel respon dan prediktor yang digunakan dalam penelitian ini disajikan dalam Tabel 3.1.

Tabel 3.1. Struktur Data Penelitian

Kab/Kota (i)	Respon		Prediktor			
	y_{1i}	y_{2i}	x_{1i}	x_{2i}	...	x_{5i}
1	y_{11}	y_{21}	x_{11}	x_{21}	...	x_{51}
2	y_{12}	y_{22}	x_{12}	x_{22}	...	x_{52}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$n = 118$	y_{1n}	y_{2n}	x_{1n}	x_{2n}	...	x_{5n}

3.4.4. Langkah-langkah Analisis

Sub-bab ini menguraikan tentang langkah-langkah penelitian yang akan dilakukan untuk menerapkan model regresi nonparametrik birespon kontinu berdasarkan estimator MARS pada pemodelan data persentase penduduk miskin dan IPM di pulau Jawa, sebagai berikut:

1. Membuat eksplorasi data dari respon (persentase penduduk miskin dan IPM) dan prediktor (pertumbuhan ekonomi, angka beban ketergantungan, persentase pengeluaran perkapita untuk makanan, tingkat pengangguran terbuka dan persentase penduduk yang pekerjaan utamanya di sektor pertanian) untuk mengetahui statistik deskriptif dari masing-masing variabel penelitian.
2. Membuat *scatter plot* data untuk mengetahui bentuk pola hubungan antara respon dan prediktor.
3. Menguji korelasi antar respon dengan uji korelasi *pearson*.
4. Menguji hubungan linier antara prediktor dengan respon dengan *Ramsey Reset Test*.
5. Mencari fungsi basis optimal dengan proses *stepwise (forward dan backward)* berdasarkan nilai GCV minimum pada setiap respon.
6. Mengestimasi matriks bobot berdasarkan fungsi basis optimal yang diperoleh dari langkah 5.
7. Mengestimasi parameter/koeffisien regresi MARS birespon kontinu berdasarkan fungsi basis optimal yang diperoleh dari langkah 5 dan estimasi matriks bobot yang diperoleh dari langkah 6.
8. Mengestimasi fungsi regresi MARS birespon kontinu berdasarkan fungsi basis optimal yang diperoleh dari langkah 5 dan estimasi parameter yang diperoleh dari langkah 7.
9. Menguji *random error* berdistribusi normal multivariat dengan Q-Q Plot.
10. Menjelaskan contoh pola hubungan antara respon dan prediktor yang dihasilkan oleh fungsi basis dalam model.
11. Melakukan validasi terhadap model yang diperoleh.

BAB 4

HASIL KAJIAN TEORITIK

Bab ini menyajikan hasil pembahasan teori untuk menjawab Tujuan 1 dan 2 yang disebutkan dalam Bab 1, yang umumnya membahas pengembangan estimator MARS dalam regresi nonparametrik birespon kontinu. Pembahasan mengenai Tujuan 1 difokuskan pada teori untuk mendapatkan estimasi fungsi regresi MARS dalam model regresi nonparametrik birespon kontinu. Pembahasan mengenai Tujuan 2 difokuskan pada kajian teori untuk mendapatkan estimasi bobot yang mengakomodasi adanya korelasi antara respon dalam model regresi MARS birespon kontinu. Metode yang digunakan untuk mendapatkan estimasi fungsi regresi dan estimasi bobot dalam penelitian ini adalah metode MLE. Hasil pembahasan teori diperoleh berdasarkan tahapan metodologi pada Bab 3, yang hasilnya dijabarkan dalam sub bab-sub bab dibawah ini.

4.1. Estimasi Fungsi Regresi dalam Model MARS Birespon Kontinu

Diberikan data berpasangan $(x_{1i}, x_{2i}, \dots, x_{pi}, y_{1i})$ dan $(x_{1i}, x_{2i}, \dots, x_{pi}, y_{2i})$ $i = 1, 2, \dots, n$, dimana x_1, x_2, \dots, x_p adalah prediktor dan y_1, y_2 adalah respon sedangkan n menunjukkan banyaknya observasi. Apabila hubungan antara prediktor dan respon dinyatakan dalam fungsi regresi f yang tidak diketahui bentuknya dan dapat didekati dengan model regresi nonparametrik birespon kontinu, maka hubungan tersebut dapat dinyatakan dalam bentuk persamaan regresi sebagai berikut:

$$y_{\ell i} = f_{\ell}(x_{1i}, x_{2i}, \dots, x_{pi}) + \varepsilon_{\ell i}; \ell = 1, 2; i = 1, 2, \dots, n. \quad (4.1)$$

Wang *et al.* (2000) dalam penelitiannya tentang birespon *smoothing* spline mengasumsikan *random error* pada persamaan (4.1), sebagai berikut :

$$E(\varepsilon_{\ell i}) = 0, E(\varepsilon_{\ell i} \varepsilon_{\ell i'}) = \begin{cases} \sigma_{\ell\ell}, & \text{jika } i = i' \\ 0, & \text{jika } i \neq i' \end{cases}, \text{ dengan } \ell = 1, 2, \text{ dan}$$

$$E(\varepsilon_{1i} \varepsilon_{2i'}) = \begin{cases} \sigma_{12}, & \text{jika } i = i' \\ 0, & \text{jika } i \neq i' \end{cases}, \text{ dengan } \sigma_{12} = \rho \sqrt{\sigma_{11} \sigma_{22}}. \quad (4.2)$$

Berdasarkan observasi ke- i untuk $\ell = 1, 2$ respon, maka model regresi nonparametrik birespon pada persamaan (4.1) dapat dijabarkan dalam persamaan sebagai berikut:

$$\begin{aligned} y_{1i} &= f_1(x_{1i}, x_{2i}, \dots, x_{pi}) + \varepsilon_{1i}, \\ y_{2i} &= f_2(x_{1i}, x_{2i}, \dots, x_{pi}) + \varepsilon_{2i}. \end{aligned} \quad (4.3)$$

Dalam bentuk matriks, persamaan (4.3) dengan n data pengamatan dapat dinyatakan dengan persamaan sebagai berikut:

$$\underline{y} = \underline{f}(\underline{x}) + \underline{\varepsilon}, \quad (4.4)$$

dimana :

$$\begin{aligned} \underline{y} &= (\underline{y}'_1, \underline{y}'_2)' = (y_{11}, y_{12}, \dots, y_{1n}, y_{21}, y_{22}, \dots, y_{2n})', \\ \underline{f}(\underline{x}) &= (\underline{f}'_1(\underline{x}), \underline{f}'_2(\underline{x}))', \\ \underline{f}_1(\underline{x}) &= [f_1(x_{11}, x_{21}, \dots, x_{p1}), f_1(x_{12}, x_{22}, \dots, x_{p2}), \dots, f_1(x_{1n}, x_{2n}, \dots, x_{pn})]', \\ \underline{f}_2(\underline{x}) &= [f_2(x_{11}, x_{21}, \dots, x_{p1}), f_2(x_{12}, x_{22}, \dots, x_{p2}), \dots, f_2(x_{1n}, x_{2n}, \dots, x_{pn})]', \\ \underline{\varepsilon} &= (\underline{\varepsilon}'_1, \underline{\varepsilon}'_2)' = (\varepsilon_{11}, \varepsilon_{12}, \dots, \varepsilon_{1n}, \varepsilon_{21}, \varepsilon_{22}, \dots, \varepsilon_{2n})'. \end{aligned}$$

Random error $\underline{\varepsilon}$ dalam model regresi nonparametrik birespon pada persamaan (4.4) memiliki distribusi sebagai berikut (Wang *et al.*, 2000):

$$\underline{\varepsilon} \sim N(0, \mathbf{W}). \quad (4.5)$$

Berdasarkan asumsi $\underline{\varepsilon}$ pada persamaan (4.5), maka matriks kovarians *error* \mathbf{W} dalam model MARS birespon kontinu dapat diuraikan dengan persamaan sebagai berikut:

$$\begin{aligned} \mathbf{W} &= \text{Var}(\underline{\varepsilon}) \\ &= E \left[(\underline{\varepsilon} - E(\underline{\varepsilon}))(\underline{\varepsilon} - E(\underline{\varepsilon}))' \right] \\ &= E(\underline{\varepsilon}\underline{\varepsilon}') \\ &= E \left[(\varepsilon_{11}, \varepsilon_{12}, \dots, \varepsilon_{1n}, \varepsilon_{21}, \varepsilon_{22}, \dots, \varepsilon_{2n})' (\varepsilon_{11}, \varepsilon_{12}, \dots, \varepsilon_{1n}, \varepsilon_{21}, \varepsilon_{22}, \dots, \varepsilon_{2n}) \right] \end{aligned}$$

$$= E \left(\begin{array}{cccc|cccc} \varepsilon_{11}^2 & \varepsilon_{11}\varepsilon_{12} & \cdots & \varepsilon_{11}\varepsilon_{1n} & \varepsilon_{11}\varepsilon_{21} & \varepsilon_{11}\varepsilon_{22} & \cdots & \varepsilon_{11}\varepsilon_{2n} \\ \varepsilon_{12}\varepsilon_{11} & \varepsilon_{12}^2 & \cdots & \varepsilon_{12}\varepsilon_{1n} & \varepsilon_{12}\varepsilon_{21} & \varepsilon_{12}\varepsilon_{22} & \cdots & \varepsilon_{12}\varepsilon_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{1n}\varepsilon_{11} & \varepsilon_{1n}\varepsilon_{12} & \cdots & \varepsilon_{1n}^2 & \varepsilon_{1n}\varepsilon_{21} & \varepsilon_{1n}\varepsilon_{22} & \cdots & \varepsilon_{1n}\varepsilon_{2n} \\ \hline \varepsilon_{21}\varepsilon_{11} & \varepsilon_{21}\varepsilon_{12} & \cdots & \varepsilon_{21}\varepsilon_{1n} & \varepsilon_{21}^2 & \varepsilon_{21}\varepsilon_{22} & \cdots & \varepsilon_{21}\varepsilon_{2n} \\ \varepsilon_{22}\varepsilon_{11} & \varepsilon_{22}\varepsilon_{12} & \cdots & \varepsilon_{22}\varepsilon_{1n} & \varepsilon_{22}\varepsilon_{21} & \varepsilon_{22}^2 & \cdots & \varepsilon_{22}\varepsilon_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{2n}\varepsilon_{11} & \varepsilon_{2n}\varepsilon_{12} & \cdots & \varepsilon_{2n}\varepsilon_{1n} & \varepsilon_{2n}\varepsilon_{21} & \varepsilon_{2n}\varepsilon_{22} & \cdots & \varepsilon_{2n}^2 \end{array} \right). \quad (4.6)$$

Matriks \mathbf{W} dalam persamaan (4.6) berorde $2n \times 2n$, dengan menggunakan asumsi pada persamaan (4.2), matriks \mathbf{W} dapat dinyatakan menjadi:

$$\mathbf{W} = \left(\begin{array}{cccc|cccc} \sigma_{11} & 0 & \cdots & 0 & \rho\sqrt{\sigma_{11}\sigma_{22}} & 0 & \cdots & 0 \\ 0 & \sigma_{11} & \cdots & 0 & 0 & \rho\sqrt{\sigma_{11}\sigma_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{11} & 0 & 0 & \cdots & \rho\sqrt{\sigma_{11}\sigma_{22}} \\ \hline \rho\sqrt{\sigma_{11}\sigma_{22}} & 0 & \cdots & 0 & \sigma_{22} & 0 & \cdots & 0 \\ 0 & \rho\sqrt{\sigma_{11}\sigma_{22}} & \cdots & 0 & 0 & \sigma_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \rho\sqrt{\sigma_{11}\sigma_{22}} & 0 & 0 & \cdots & \sigma_{22} \end{array} \right) \quad (4.7)$$

$$= \left(\begin{array}{cccc|cccc} \sigma_{11} & 0 & \cdots & 0 & \sigma_{12} & 0 & \cdots & 0 \\ 0 & \sigma_{11} & \cdots & 0 & 0 & \sigma_{12} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{11} & 0 & 0 & \cdots & \sigma_{12} \\ \hline \sigma_{12} & 0 & \cdots & 0 & \sigma_{22} & 0 & \cdots & 0 \\ 0 & \sigma_{12} & \cdots & 0 & 0 & \sigma_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{12} & 0 & 0 & \cdots & \sigma_{22} \end{array} \right) \quad (4.8)$$

$$= \left(\begin{array}{cc|cc} \sigma_{11} & & & \sigma_{12} \\ \hline & & & \\ \hline \sigma_{12} & & & \sigma_{22} \end{array} \right) \otimes \mathbf{I} \quad (4.9)$$

dengan \mathbf{I} adalah matriks identitas berorde n .

Berdasarkan persamaan regresi nonparametrik birespon pada persamaan (4.3), maka dapat diperoleh persamaan regresi MARS birespon kontinu yang disajikan dalam Lemma 4.1.

Lemma 4.1.

Jika diberikan data berpasangan $(x_{1i}, x_{2i}, \dots, x_{pi}, y_{1i})$ dan $(x_{1i}, x_{2i}, \dots, x_{pi}, y_{2i})$ $i = 1, 2, \dots, n$ dan hubungan antara prediktor $(x_{1i}, x_{2i}, \dots, x_{pi})$ dan respon (y_{1i}, y_{2i}) dapat didekati dengan fungsi regresi nonparametrik birespon kontinu seperti yang diberikan dalam persamaan (4.3), maka model persamaan regresi MARS birespon kontinu adalah sebagai berikut:

$$\underline{y} = \mathbf{B}(\underline{x}, \underline{t})\underline{\alpha} + \underline{\varepsilon}, \quad (4.10)$$

dimana:

$$\underline{y} = (y'_1, y'_2)' = (y_{11}, y_{12}, \dots, y_{1n}, y_{21}, y_{22}, \dots, y_{2n})',$$

$$\mathbf{B}(\underline{x}, \underline{t}) = \begin{pmatrix} \mathbf{B}_1(\underline{x}, \underline{t}) & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_2(\underline{x}, \underline{t}) \end{pmatrix},$$

dengan :

$$\mathbf{B}_\ell(\underline{x}, \underline{t}) = \begin{pmatrix} 1 \prod_{k_\ell=1}^{K_1} [s_{k_\ell 1}(x_{v(k_\ell, 1)1} - t_{k_\ell 1})]_+ & \dots & \prod_{k_\ell=1}^{K_{M_\ell}} [s_{k_\ell M_\ell}(x_{v(k_\ell, M_\ell)1} - t_{k_\ell M_\ell})]_+ \\ 1 \prod_{k_\ell=1}^{K_1} [s_{k_\ell 1}(x_{v(k_\ell, 1)2} - t_{k_\ell 1})]_+ & \dots & \prod_{k_\ell=1}^{K_{M_\ell}} [s_{k_\ell M_\ell}(x_{v(k_\ell, M_\ell)2} - t_{k_\ell M_\ell})]_+ \\ \vdots & \ddots & \vdots \\ 1 \prod_{k_\ell=1}^{K_1} [s_{k_\ell 1}(x_{v(k_\ell, 1)n} - t_{k_\ell 1})]_+ & \dots & \prod_{k_\ell=1}^{K_{M_\ell}} [s_{k_\ell M_\ell}(x_{v(k_\ell, M_\ell)n} - t_{k_\ell M_\ell})]_+ \end{pmatrix}, \quad \ell = 1, 2,$$

$$\underline{\alpha} = (\alpha'_1, \alpha'_2)' = (\alpha_{11}, \alpha_{12}, \dots, \alpha_{1M_1}, \alpha_{21}, \alpha_{22}, \dots, \alpha_{2M_2})',$$

$$\underline{\varepsilon} = (\varepsilon'_1, \varepsilon'_2)' = (\varepsilon_{11}, \varepsilon_{12}, \dots, \varepsilon_{1n}, \varepsilon_{21}, \varepsilon_{22}, \dots, \varepsilon_{2n})',$$

dengan $\underline{\varepsilon} \sim N(0, \mathbf{W})$.

Bukti :

Fungsi regresi f_1 dan f_2 dalam persamaan (4.3) merupakan fungsi regresi nonparametrik birespon kontinu yang diasumsikan tidak diketahui bentuknya. Karena fungsi-fungsi tersebut dihipotesiskan dengan fungsi regresi MARS seperti yang dinyatakan dalam persamaan (2.12), maka fungsi regresi MARS birespon kontinu dapat dinyatakan untuk fungsi regresi f_1 sebagai berikut:

$$f_1(x_{1i}, x_{2i}, \dots, x_{pi}) = \alpha_{10} + \sum_{m_1=1}^{M_1} \alpha_{1m_1} \prod_{k_1=1}^{K_{m_1}} [s_{k_1 m_1} (x_{v(k_1, m_1)i} - t_{k_1 m_1})]_+$$

dengan:

$$x_{v(k_1, m_1)} \in \{x_j\}_{j=1}^p, \quad t_{k_1 m_1} \in \{x_{v(k_1, m_1)i}\}_{i=1}^n, \quad m_1 = 1, 2, \dots, M_1,$$

$$\text{jika } s_{k_1 m_1} = +1, \text{ maka } (x_{v(k_1, m_1)i} - t_{k_1 m_1})_+ = \begin{cases} x_{v(k_1, m_1)i} - t_{k_1 m_1}, & \text{jika } x_{v(k_1, m_1)i} > t_{k_1 m_1} \\ 0 & , \text{sebaliknya} \end{cases},$$

$$\text{jika } s_{k_1 m_1} = -1, \text{ maka } -(x_{v(k_1, m_1)i} - t_{k_1 m_1})_+ = \begin{cases} t_{k_1 m_1} - x_{v(k_1, m_1)i}, & \text{jika } t_{k_1 m_1} > x_{v(k_1, m_1)i} \\ 0 & , \text{sebaliknya} \end{cases},$$

dan fungsi regresi f_2 sebagai berikut:

$$f_2(x_{1i}, x_{2i}, \dots, x_{pi}) = \alpha_{20} + \sum_{m_2=1}^{M_2} \alpha_{2m_2} \prod_{k_2=1}^{K_{m_2}} [s_{k_2 m_2} (x_{v(k_2, m_2)i} - t_{k_2 m_2})]_+$$

dengan:

$$x_{v(k_2, m_2)} \in \{x_j\}_{j=1}^p, \quad t_{k_2 m_2} \in \{x_{v(k_2, m_2)i}\}_{i=1}^n, \quad m_2 = 1, 2, \dots, M_2,$$

$$\text{jika } s_{k_2 m_2} = +1, \text{ maka } (x_{v(k_2, m_2)i} - t_{k_2 m_2})_+ = \begin{cases} x_{v(k_2, m_2)i} - t_{k_2 m_2}, & \text{jika } x_{v(k_2, m_2)i} > t_{k_2 m_2} \\ 0 & , \text{sebaliknya} \end{cases},$$

$$\text{jika } s_{k_2 m_2} = -1, \text{ maka } -(x_{v(k_2, m_2)i} - t_{k_2 m_2})_+ = \begin{cases} t_{k_2 m_2} - x_{v(k_2, m_2)i}, & \text{jika } t_{k_2 m_2} > x_{v(k_2, m_2)i} \\ 0 & , \text{sebaliknya} \end{cases}.$$

(4.11)

Fungsi regresi MARS birespon kontinu dalam persamaan (4.11) dapat disederhanakan untuk fungsi regresi f_1 dan f_2 sebagai berikut:

$$\begin{aligned} f_1(x_{1i}, x_{2i}, \dots, x_{pi}) &= \alpha_{10} + \sum_{m_1=1}^{M_1} \alpha_{1m_1} B_{m_1 i}(x, t), \\ f_2(x_{1i}, x_{2i}, \dots, x_{pi}) &= \alpha_{20} + \sum_{m_2=1}^{M_2} \alpha_{2m_2} B_{m_2 i}(x, t). \end{aligned} \quad (4.12)$$

Dengan $B_{m_1i}(\underline{x}, \underline{t})$ dan $B_{m_2i}(\underline{x}, \underline{t})$ pada persamaan (4.12) masing-masing adalah fungsi basis untuk $f_1(x_{1i}, x_{2i}, \dots, x_{pi})$ dan fungsi basis untuk $f_2(x_{1i}, x_{2i}, \dots, x_{pi})$ dalam model MARS, yang dapat dinyatakan dalam bentuk persamaan sebagai berikut :

$$\begin{aligned} B_{m_1i}(\underline{x}, \underline{t}) &= \prod_{k_1=1}^{K_{m_1}} \left[s_{k_1 m_1} (x_{v(k_1, m_1)i} - t_{k_1 m_1}) \right]_+, \\ B_{m_2i}(\underline{x}, \underline{t}) &= \prod_{k_2=1}^{K_{m_2}} \left[s_{k_2 m_2} (x_{v(k_2, m_2)i} - t_{k_2 m_2}) \right]_+. \end{aligned} \quad (4.13)$$

Apabila diuraikan untuk setiap observasi ($i=1, 2, \dots, n$), maka fungsi regresi MARS birespon kontinu dalam persamaan (4.11) dapat diuraikan untuk fungsi regresi f_1 sebagai berikut:

$$\begin{aligned} f_1(x_{11}, x_{21}, \dots, x_{p1}) &= \alpha_{10} + \sum_{m_1=1}^{M_1} \alpha_{1m_1} \prod_{k_1=1}^{K_{m_1}} \left[s_{k_1 m_1} (x_{v(k_1, m_1)1} - t_{k_1 m_1}) \right]_+, \\ f_1(x_{12}, x_{22}, \dots, x_{p2}) &= \alpha_{10} + \sum_{m_1=1}^{M_1} \alpha_{1m_1} \prod_{k_1=1}^{K_{m_1}} \left[s_{k_1 m_1} (x_{v(k_1, m_1)2} - t_{k_1 m_1}) \right]_+, \\ &\vdots \\ f_1(x_{1n}, x_{2n}, \dots, x_{pn}) &= \alpha_{10} + \sum_{m_1=1}^{M_1} \alpha_{1m_1} \prod_{k_1=1}^{K_{m_1}} \left[s_{k_1 m_1} (x_{v(k_1, m_1)n} - t_{k_1 m_1}) \right]_+, \end{aligned}$$

dan fungsi regresi f_2 sebagai berikut:

$$\begin{aligned} f_2(x_{11}, x_{21}, \dots, x_{p1}) &= \alpha_{20} + \sum_{m_2=1}^{M_2} \alpha_{2m_2} \prod_{k_2=1}^{K_{m_2}} \left[s_{k_2 m_2} (x_{v(k_2, m_2)1} - t_{k_2 m_2}) \right]_+, \\ f_2(x_{12}, x_{22}, \dots, x_{p2}) &= \alpha_{20} + \sum_{m_2=1}^{M_2} \alpha_{2m_2} \prod_{k_2=1}^{K_{m_2}} \left[s_{k_2 m_2} (x_{v(k_2, m_2)2} - t_{k_2 m_2}) \right]_+, \\ &\vdots \\ f_2(x_{1n}, x_{2n}, \dots, x_{pn}) &= \alpha_{20} + \sum_{m_2=1}^{M_2} \alpha_{2m_2} \prod_{k_2=1}^{K_{m_2}} \left[s_{k_2 m_2} (x_{v(k_2, m_2)n} - t_{k_2 m_2}) \right]_+. \end{aligned} \quad (4.14)$$

Berdasarkan banyaknya fungsi basis ($m_1 = 1, 2, \dots, M_1$ dan $m_2 = 1, 2, \dots, M_2$), maka fungsi regresi MARS birespon kontinu dalam persamaan (4.14) dapat diuraikan untuk fungsi regresi f_1 sebagai berikut:

$$\begin{aligned}
f_1(x_{11}, \dots, x_{p1}) &= \alpha_{10} + \alpha_{11} \prod_{k_1=1}^{K_1} [s_{k_11}(x_{v(k_1,1)1} - t_{k_11})]_+ + \dots + \alpha_{1M_1} \prod_{k_1=1}^{K_{M_1}} [s_{k_1M_1}(x_{v(k_1,M_1)1} - t_{k_1M_1})]_+, \\
f_1(x_{12}, \dots, x_{p2}) &= \alpha_{10} + \alpha_{11} \prod_{k_1=1}^{K_1} [s_{k_11}(x_{v(k_1,1)2} - t_{k_11})]_+ + \dots + \alpha_{1M_1} \prod_{k_1=1}^{K_{M_1}} [s_{k_1M_1}(x_{v(k_1,M_1)2} - t_{k_1M_1})]_+, \\
&\vdots \\
f_1(x_{1n}, \dots, x_{pn}) &= \alpha_{10} + \alpha_{11} \prod_{k_1=1}^{K_1} [s_{k_11}(x_{v(k_1,1)n} - t_{k_11})]_+ + \dots + \alpha_{1M_1} \prod_{k_1=1}^{K_{M_1}} [s_{k_1M_1}(x_{v(k_1,M_1)n} - t_{k_1M_1})]_+,
\end{aligned}$$

dan fungsi regresi f_2 sebagai berikut:

$$\begin{aligned}
f_2(x_{11}, \dots, x_{p1}) &= \alpha_{20} + \alpha_{21} \prod_{k_2=1}^{K_1} [s_{k_21}(x_{v(k_2,1)1} - t_{k_21})]_+ + \dots + \alpha_{2M_2} \prod_{k_2=1}^{K_{M_2}} [s_{k_2M_2}(x_{v(k_2,M_2)1} - t_{k_2M_2})]_+, \\
f_2(x_{12}, \dots, x_{p2}) &= \alpha_{20} + \alpha_{21} \prod_{k_2=1}^{K_1} [s_{k_21}(x_{v(k_2,1)2} - t_{k_21})]_+ + \dots + \alpha_{2M_2} \prod_{k_2=1}^{K_{M_2}} [s_{k_2M_2}(x_{v(k_2,M_2)2} - t_{k_2M_2})]_+, \\
&\vdots \\
f_2(x_{1n}, \dots, x_{pn}) &= \alpha_{20} + \alpha_{21} \prod_{k_2=1}^{K_1} [s_{k_21}(x_{v(k_2,1)n} - t_{k_21})]_+ + \dots + \alpha_{2M_2} \prod_{k_2=1}^{K_{M_2}} [s_{k_2M_2}(x_{v(k_2,M_2)n} - t_{k_2M_2})]_+.
\end{aligned} \tag{4.15}$$

Fungsi regresi MARS birespon kontinu dalam persamaan (4.15), dapat disederhanakan dalam bentuk matriks sebagai berikut:

$$\tilde{f}(\tilde{x}) = \mathbf{B}(\tilde{x}, \tilde{t})\tilde{\alpha}, \tag{4.16}$$

dimana $\tilde{f}(\tilde{x})$ adalah fungsi regresi MARS birespon kontinu yang memuat $\tilde{f}_1(\tilde{x})$

dan $\tilde{f}_2(\tilde{x})$, yang dinyatakan dalam bentuk vektor sebagai berikut:

$$\tilde{f}(\tilde{x}) = \begin{pmatrix} \tilde{f}_1(\tilde{x}) \\ \tilde{f}_2(\tilde{x}) \end{pmatrix} = \begin{pmatrix} f_1(x_{11}, x_{21}, \dots, x_{p1}) \\ f_1(x_{12}, x_{22}, \dots, x_{p2}) \\ \vdots \\ f_1(x_{1n}, x_{2n}, \dots, x_{pn}) \\ f_2(x_{11}, x_{21}, \dots, x_{p1}) \\ f_2(x_{12}, x_{22}, \dots, x_{p2}) \\ \vdots \\ f_2(x_{1n}, x_{2n}, \dots, x_{pn}) \end{pmatrix},$$

$\mathbf{B}(\underline{x}, \underline{t})$ adalah matriks diagonal yang memuat fungsi basis $\mathbf{B}_1(\underline{x}, \underline{t})$ untuk $f_1(\underline{x})$ dan fungsi basis $\mathbf{B}_2(\underline{x}, \underline{t})$ untuk $f_2(\underline{x})$ sebagai berikut:

$$\mathbf{B}(\underline{x}, \underline{t}) = \begin{pmatrix} \mathbf{B}_1(\underline{x}, \underline{t}) & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_2(\underline{x}, \underline{t}) \end{pmatrix}$$

dengan:

$$\mathbf{B}_1(\underline{x}, \underline{t}) = \begin{pmatrix} 1 \prod_{k_1=1}^{K_1} [s_{k_1 1}(x_{v(k_1,1)1} - t_{k_1 1})]_+ & \cdots & \prod_{k_1=1}^{K_{M_1}} [s_{k_1 M_1}(x_{v(k_1, M_1)1} - t_{k_1 M_1})]_+ \\ 1 \prod_{k_1=1}^{K_1} [s_{k_1 1}(x_{v(k_1,1)2} - t_{k_1 1})]_+ & \cdots & \prod_{k_1=1}^{K_{M_1}} [s_{k_1 M_1}(x_{v(k_1, M_1)2} - t_{k_1 M_1})]_+ \\ \vdots & \vdots & \vdots \\ 1 \prod_{k_1=1}^{K_1} [s_{k_1 1}(x_{v(k_1,1)n} - t_{k_1 1})]_+ & \cdots & \prod_{k_1=1}^{K_{M_1}} [s_{k_1 M_1}(x_{v(k_1, M_1)n} - t_{k_1 M_1})]_+ \end{pmatrix},$$

$$\mathbf{B}_2(\underline{x}, \underline{t}) = \begin{pmatrix} 1 \prod_{k_2=1}^{K_1} [s_{k_2 1}(x_{v(k_2,1)1} - t_{k_2 1})]_+ & \cdots & \prod_{k_2=1}^{K_{M_2}} [s_{k_2 M_2}(x_{v(k_2, M_2)1} - t_{k_2 M_2})]_+ \\ 1 \prod_{k_2=1}^{K_1} [s_{k_2 1}(x_{v(k_2,1)2} - t_{k_2 1})]_+ & \cdots & \prod_{k_2=1}^{K_{M_2}} [s_{k_2 M_2}(x_{v(k_2, M_2)2} - t_{k_2 M_2})]_+ \\ \vdots & \vdots & \vdots \\ 1 \prod_{k_2=1}^{K_1} [s_{k_2 1}(x_{v(k_2,1)n} - t_{k_2 1})]_+ & \cdots & \prod_{k_2=1}^{K_{M_2}} [s_{k_2 M_2}(x_{v(k_2, M_2)n} - t_{k_2 M_2})]_+ \end{pmatrix},$$

sedangkan $\underline{\alpha}$ adalah parameter atau koefisien regresi MARS birespon kontinu yang memuat parameter α_1 untuk $\mathbf{B}_1(\underline{x}, \underline{t})$ dan parameter α_2 untuk $\mathbf{B}_2(\underline{x}, \underline{t})$ yang dinyatakan dalam bentuk vektor sebagai berikut :

$$\underline{\alpha} = \begin{pmatrix} \underline{\alpha}_1 \\ \underline{\alpha}_2 \end{pmatrix} = \begin{pmatrix} \alpha_{10} \\ \alpha_{11} \\ \vdots \\ \alpha_{1M_1} \\ \alpha_{20} \\ \alpha_{21} \\ \vdots \\ \alpha_{2M_2} \end{pmatrix}.$$

Bentuk persamaan regresi MARS birespon kontinu, selanjutnya diperoleh dengan mensubstitusikan persamaan (4.16) ke dalam persamaan (4.4), sehingga diperoleh:

$$\begin{aligned}\tilde{y} &= \tilde{f}(\tilde{x}) + \varepsilon \\ &= \mathbf{B}(\tilde{x}, \tilde{t})\alpha + \varepsilon.\end{aligned}$$

■

Random error ε dalam persamaan regresi MARS birespon kontinu berdistribusi normal dengan *mean* 0 dan matriks kovarians *error* \mathbf{W} . Matriks \mathbf{W} berfungsi sebagai matriks bobot yang mengakomodasi adanya korelasi antar respon pada saat estimasi parameter model. Untuk mendapatkan estimasi parameter model, matriks \mathbf{W} diasumsikan diketahui nilainya (*fixed variable*). Hasil estimasi dari fungsi regresi MARS birespon kontinu diuraikan dalam Teorema 4.1.

Teorema 4.1.

Jika persamaan regresi MARS birespon kontinu diberikan oleh Lemma 4.1, dengan *random error* ε diasumsikan berdistribusi normal dengan *mean* 0 dan matriks kovarians *error* \mathbf{W} , dimana matriks \mathbf{W} diasumsikan diketahui, maka estimasi fungsi regresi MARS birespon kontinu $\tilde{f}(\tilde{x})$ dinyatakan dengan persamaan sebagai berikut :

$$\hat{\tilde{f}}(\tilde{x}) = \mathbf{H}(\cdot)\tilde{y}, \quad (4.17)$$

dimana:

$$\mathbf{H}(\cdot) = \mathbf{B}(\tilde{x}, \tilde{t}) \left(\mathbf{B}'(\tilde{x}, \tilde{t}) \mathbf{W}^{-1} \mathbf{B}(\tilde{x}, \tilde{t}) \right)^{-1} \mathbf{B}'(\tilde{x}, \tilde{t}) \mathbf{W}^{-1}.$$

Bukti:

Karena *random error* ε dalam persamaan regresi MARS birespon kontinu diasumsikan berdistribusi normal dengan *mean* 0 (vektor berorde $2n \times 1$) dan matriks kovarians *error* \mathbf{W} (matriks berorde $2n \times 2n$), dimana matriks \mathbf{W}

diasumsikan diketahui, maka fungsi probabilitas bersama dari $\boldsymbol{\varepsilon} = (\varepsilon_{11}, \varepsilon_{12}, \dots, \varepsilon_{1n}, \varepsilon_{21}, \varepsilon_{22}, \dots, \varepsilon_{2n})'$ diberikan oleh persamaan berikut ini:

$$g(\boldsymbol{\varepsilon}) = \frac{1}{(2\pi)^{2n/2} |\mathbf{W}|^{1/2}} \exp \left[-\frac{1}{2} \boldsymbol{\varepsilon}' \mathbf{W}^{-1} \boldsymbol{\varepsilon} \right]. \quad (4.18)$$

Karena $\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{f}(\mathbf{x})$, dengan $\mathbf{f}(\mathbf{x})$ dihampiri dengan fungsi regresi MARS birespon kontinu yang diberikan oleh Lemma 4.1, sehingga $\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{B}(\mathbf{x}, \mathbf{t})\boldsymbol{\alpha}$, maka fungsi *likelihood* dari persamaan (4.18) dinyatakan sebagai berikut:

$$\begin{aligned} L(\boldsymbol{\alpha} | \mathbf{W}, \mathbf{y}) &= \frac{1}{(2\pi)^{2n/2} |\mathbf{W}|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{y} - \mathbf{B}(\mathbf{x}, \mathbf{t})\boldsymbol{\alpha})' \mathbf{W}^{-1} (\mathbf{y} - \mathbf{B}(\mathbf{x}, \mathbf{t})\boldsymbol{\alpha}) \right] \\ &= (2\pi)^{-n} |\mathbf{W}|^{-1/2} \exp \left[-\frac{1}{2} (\mathbf{y} - \mathbf{B}(\mathbf{x}, \mathbf{t})\boldsymbol{\alpha})' \mathbf{W}^{-1} (\mathbf{y} - \mathbf{B}(\mathbf{x}, \mathbf{t})\boldsymbol{\alpha}) \right]. \end{aligned} \quad (4.19)$$

Dengan mengambil transformasi \ln , maka fungsi *log likelihood* dalam persamaan (4.19), dapat dinyatakan dengan persamaan sebagai berikut:

$$\begin{aligned} \ln L(\boldsymbol{\alpha} | \mathbf{W}, \mathbf{y}) &= -n \ln(2\pi) - \frac{1}{2} \ln(|\mathbf{W}|) - \frac{1}{2} (\mathbf{y} - \mathbf{B}(\mathbf{x}, \mathbf{t})\boldsymbol{\alpha})' \mathbf{W}^{-1} (\mathbf{y} - \mathbf{B}(\mathbf{x}, \mathbf{t})\boldsymbol{\alpha}) \\ &= -n \ln(2\pi) - \frac{1}{2} \ln(|\mathbf{W}|) - \frac{1}{2} (\mathbf{y} - \mathbf{B}(\mathbf{x}, \mathbf{t})\boldsymbol{\alpha})' (\mathbf{W}^{-1} \mathbf{y} - \mathbf{W}^{-1} \mathbf{B}(\mathbf{x}, \mathbf{t})\boldsymbol{\alpha}) \\ &= -n \ln(2\pi) - \frac{1}{2} \ln(|\mathbf{W}|) - \frac{1}{2} (\mathbf{y}' - \boldsymbol{\alpha}' \mathbf{B}'(\mathbf{x}, \mathbf{t})) (\mathbf{W}^{-1} \mathbf{y} - \mathbf{W}^{-1} \mathbf{B}(\mathbf{x}, \mathbf{t})\boldsymbol{\alpha}) \\ &= -n \ln(2\pi) - \frac{1}{2} \ln(|\mathbf{W}|) - \\ &\quad \frac{1}{2} (\mathbf{y}' \mathbf{W}^{-1} \mathbf{y} - \mathbf{y}' \mathbf{W}^{-1} \mathbf{B}(\mathbf{x}, \mathbf{t})\boldsymbol{\alpha} - \boldsymbol{\alpha}' \mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{y} + \boldsymbol{\alpha}' \mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{B}(\mathbf{x}, \mathbf{t})\boldsymbol{\alpha}). \end{aligned} \quad (4.20)$$

Nilai $\mathbf{y}' \mathbf{W}^{-1} \mathbf{B}(\mathbf{x}, \mathbf{t})\boldsymbol{\alpha}$ adalah skalar sehingga nilainya sama dengan $\boldsymbol{\alpha}' \mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{y}$, dengan demikian maka fungsi *log likelihood* dalam persamaan (4.20) dapat dinyatakan kembali menjadi:

$$\begin{aligned} \ln L(\boldsymbol{\alpha} | \mathbf{W}, \mathbf{y}) &= -n \ln(2\pi) - \frac{1}{2} \ln(|\mathbf{W}|) - \\ &\quad \frac{1}{2} (\mathbf{y}' \mathbf{W}^{-1} \mathbf{y} - \boldsymbol{\alpha}' \mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{y} - \boldsymbol{\alpha}' \mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{y} + \boldsymbol{\alpha}' \mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{B}(\mathbf{x}, \mathbf{t})\boldsymbol{\alpha}) \end{aligned}$$

$$\begin{aligned}
&= -n \ln(2\pi) - \frac{1}{2} \ln(|\mathbf{W}|) - \frac{1}{2} \left(\mathbf{y}' \mathbf{W}^{-1} \mathbf{y} - 2\alpha' \mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{y} + \alpha' \mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{B}(\mathbf{x}, \mathbf{t}) \alpha \right) \\
&= -n \ln(2\pi) - \frac{1}{2} \ln(|\mathbf{W}|) - \frac{1}{2} \mathbf{y}' \mathbf{W}^{-1} \mathbf{y} + \alpha' \mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{y} - \frac{1}{2} \alpha' \mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{B}(\mathbf{x}, \mathbf{t}) \alpha.
\end{aligned} \tag{4.21}$$

Estimasi parameter α dalam persamaan (4.21) dapat diperoleh dengan menurunkan secara parsial $\ln L(\alpha | \mathbf{W}, \mathbf{y})$ terhadap α menggunakan proposisi 2.8.1 dan 2.8.2 dalam Searle *et al.* (2006) dan menyamakan hasilnya dengan 0, sebagai berikut:

$$\frac{\partial}{\partial \alpha} \ln L(\hat{\alpha} | \mathbf{W}, \mathbf{y}) = \mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{y} - \mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{B}(\mathbf{x}, \mathbf{t}) \hat{\alpha} = 0.$$

Akibatnya diperoleh persamaan:

$$\mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{B}(\mathbf{x}, \mathbf{t}) \hat{\alpha} = \mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{y}. \tag{4.22}$$

Karena $\mathbf{B}(\mathbf{x}, \mathbf{t})$ adalah matriks *full rank* dan \mathbf{W} adalah matriks *positive definite*, maka $\mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{B}(\mathbf{x}, \mathbf{t})$ adalah matriks *nonsingular*, sehingga dengan mengalikan kedua ruas persamaan (4.22) dengan $(\mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{B}(\mathbf{x}, \mathbf{t}))^{-1}$, diperoleh estimasi parameter untuk α sebagai berikut:

$$\hat{\alpha} = (\mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{B}(\mathbf{x}, \mathbf{t}))^{-1} \mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{y}. \tag{4.23}$$

Selanjutnya dengan mensubstitusikan persamaan (4.23) ke persamaan (4.16), maka diperoleh estimasi fungsi regresi nonparametrik birespon kontinu $\hat{f}(\mathbf{x})$ berdasarkan fungsi MARS sebagai berikut :

$$\begin{aligned}
\hat{f}(\mathbf{x}) &= \begin{bmatrix} \hat{f}_1(\mathbf{x}) \\ \hat{f}_2(\mathbf{x}) \end{bmatrix} = \mathbf{B}(\mathbf{x}, \mathbf{t}) \hat{\alpha} \\
&= \mathbf{B}(\mathbf{x}, \mathbf{t}) (\mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{B}(\mathbf{x}, \mathbf{t}))^{-1} \mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{y} \\
&= \mathbf{H}(\cdot) \mathbf{y},
\end{aligned} \tag{4.24}$$

dimana :

$$\mathbf{H}(\cdot) = \mathbf{B}(\mathbf{x}, \mathbf{t}) (\mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1} \mathbf{B}(\mathbf{x}, \mathbf{t}))^{-1} \mathbf{B}'(\mathbf{x}, \mathbf{t}) \mathbf{W}^{-1}. \quad \blacksquare$$

Hasil estimasi fungsi regresi MARS birespon kontinu $\hat{f}(\underline{x})$ pada persamaan (4.24) menunjukkan bahwa estimator tersebut memiliki sifat linier dalam observasi \underline{y} .

4.2. Estimasi Matriks Bobot

Estimasi fungsi regresi MARS birespon kontinu memuat matriks kovarians *error* \mathbf{W} yang berfungsi sebagai matriks bobot dalam mengakomodasi adanya korelasi antara respon. Pada umumnya matriks \mathbf{W} tidak diketahui sehingga diestimasi dari data. Hasil estimasi dari matriks \mathbf{W} diuraikan dalam Teorema 4.2.

Teorema 4.2.

Jika persamaan regresi MARS birespon kontinu diberikan oleh Lemma 4.1 dengan *random error* ε diasumsikan berdistribusi normal dengan *mean* 0 dan matriks kovarians *error* \mathbf{W} , dengan elemen matriks \mathbf{W} diberikan oleh persamaan (4.9), maka estimasi dari matriks \mathbf{W} adalah sebagai berikut:

$$\hat{\mathbf{W}} = \begin{pmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} \\ \hat{\sigma}_{12} & \hat{\sigma}_{22} \end{pmatrix} \otimes \mathbf{I}, \quad (4.25)$$

dengan:

$$\hat{\sigma}_{11} = \frac{\underline{y}_1' (\mathbf{I} - \mathbf{H}_1(\cdot))' (\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1}{n},$$

$$\hat{\sigma}_{22} = \frac{\underline{y}_2' (\mathbf{I} - \mathbf{H}_2(\cdot))' (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2}{n},$$

$$\hat{\sigma}_{12} = \frac{\underline{y}_1' (\mathbf{I} - \mathbf{H}_1(\cdot))' (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2}{n},$$

$$\mathbf{H}_1(\cdot) = \mathbf{B}_1(\underline{x}, \underline{t}) (\mathbf{B}_1'(\underline{x}, \underline{t}) \mathbf{B}_1(\underline{x}, \underline{t}))^{-1} \mathbf{B}_1'(\underline{x}, \underline{t}),$$

$$\mathbf{H}_2(\cdot) = \mathbf{B}_2(\underline{x}, \underline{t}) (\mathbf{B}_2'(\underline{x}, \underline{t}) \mathbf{B}_2(\underline{x}, \underline{t}))^{-1} \mathbf{B}_2'(\underline{x}, \underline{t}).$$

Bukti :

Karena diasumsikan $\underline{\varepsilon} \sim N(0, \mathbf{W})$, dengan $\underline{\varepsilon}$ adalah vektor berorde $2n \times 1$ sedangkan \mathbf{W} adalah matriks berorde $2n \times 2n$, maka fungsi probabilitas bersama dari $\underline{\varepsilon} = (\varepsilon_{11}, \varepsilon_{12}, \dots, \varepsilon_{1n}, \varepsilon_{21}, \varepsilon_{22}, \dots, \varepsilon_{2n})'$ diberikan oleh persamaan berikut ini:

$$g(\underline{\varepsilon}) = (2\pi)^{-2n/2} |\mathbf{W}|^{-1/2} \exp \left[-\frac{1}{2} \underline{\varepsilon}' \mathbf{W}^{-1} \underline{\varepsilon} \right]. \quad (4.26)$$

Karena $\underline{\varepsilon} = \underline{y} - \underline{f}(\underline{x})$, dimana $\underline{f}(\underline{x}) = \mathbf{B}(\underline{x}, \underline{t}) \underline{\alpha}$ adalah fungsi MARS yang diberikan oleh persamaan (4.16), maka fungsi *likelihood* dari persamaan (4.26) adalah sebagai berikut:

$$L(\mathbf{W} | \underline{\alpha}, \underline{y}) = (2\pi)^{-2n/2} |\mathbf{W}|^{-1/2} \exp \left[-\frac{1}{2} (\underline{y} - \mathbf{B}(\underline{x}, \underline{t}) \underline{\alpha})' \mathbf{W}^{-1} (\underline{y} - \mathbf{B}(\underline{x}, \underline{t}) \underline{\alpha}) \right]. \quad (4.27)$$

Karena $\underline{y} - \mathbf{B}(\underline{x}, \underline{t}) \underline{\alpha}$ dalam persamaan regresi MARS birespon kontinu dapat diuraikan dalam bentuk:

$$\begin{pmatrix} y_1 - \mathbf{B}_1(\underline{x}, \underline{t}) \alpha_1 \\ y_2 - \mathbf{B}_2(\underline{x}, \underline{t}) \alpha_2 \end{pmatrix}$$

maka fungsi *likelihood* dalam persamaan (4.27) dapat dinyatakan kembali menjadi:

$$L(\mathbf{W} | \underline{\alpha}, \underline{y}) = (2\pi)^{-n} |\mathbf{W}|^{-1/2} \exp \left[-\frac{1}{2} \begin{pmatrix} y_1 - \mathbf{B}_1(\underline{x}, \underline{t}) \alpha_1 \\ y_2 - \mathbf{B}_2(\underline{x}, \underline{t}) \alpha_2 \end{pmatrix}' \mathbf{W}^{-1} \begin{pmatrix} y_1 - \mathbf{B}_1(\underline{x}, \underline{t}) \alpha_1 \\ y_2 - \mathbf{B}_2(\underline{x}, \underline{t}) \alpha_2 \end{pmatrix} \right], \quad (4.28)$$

dimana α_1 dan α_2 dalam persamaan (4.28) diperoleh dari optimasi OLS yang tidak melibatkan matriks kovarians *error* \mathbf{W} seperti dinyatakan dalam persamaan sebagai berikut:

$$\alpha_1 = (\mathbf{B}_1'(\underline{x}, \underline{t}) \mathbf{B}_1(\underline{x}, \underline{t}))^{-1} \mathbf{B}_1'(\underline{x}, \underline{t}) y_1,$$

$$\alpha_2 = (\mathbf{B}_2'(\underline{x}, \underline{t}) \mathbf{B}_2(\underline{x}, \underline{t}))^{-1} \mathbf{B}_2'(\underline{x}, \underline{t}) y_2,$$

sehingga:

$$\mathbf{B}_1(\underline{x}, \underline{t}) \alpha_1 = \mathbf{B}_1(\underline{x}, \underline{t}) (\mathbf{B}_1'(\underline{x}, \underline{t}) \mathbf{B}_1(\underline{x}, \underline{t}))^{-1} \mathbf{B}_1'(\underline{x}, \underline{t}) y_1 = \mathbf{H}_1(\cdot) y_1,$$

$$\mathbf{B}_2(\underline{x}, \underline{t}) \alpha_2 = \mathbf{B}_2(\underline{x}, \underline{t}) (\mathbf{B}_2'(\underline{x}, \underline{t}) \mathbf{B}_2(\underline{x}, \underline{t}))^{-1} \mathbf{B}_2'(\underline{x}, \underline{t}) y_2 = \mathbf{H}_2(\cdot) y_2, \quad (4.29)$$

dengan:

$$\mathbf{H}_1(\cdot) = \mathbf{B}_1(\underline{x}, \underline{t}) (\mathbf{B}_1'(\underline{x}, \underline{t}) \mathbf{B}_1(\underline{x}, \underline{t}))^{-1} \mathbf{B}_1'(\underline{x}, \underline{t}),$$

$$\mathbf{H}_2(\cdot) = \mathbf{B}_2(\underline{x}, \underline{t}) (\mathbf{B}_2'(\underline{x}, \underline{t}) \mathbf{B}_2(\underline{x}, \underline{t}))^{-1} \mathbf{B}_2'(\underline{x}, \underline{t}).$$

Dengan mensubstitusikan persamaan (4.29) ke dalam persamaan (4.28) maka diperoleh:

$$\begin{aligned} L(\mathbf{W} | \underline{\alpha}, \underline{y}) &= (2\pi)^{-n} |\mathbf{W}|^{-1/2} \exp \left[-\frac{1}{2} \begin{pmatrix} \underline{y}_1 - \mathbf{H}_1(\cdot) \underline{y}_1 \\ \underline{y}_2 - \mathbf{H}_2(\cdot) \underline{y}_2 \end{pmatrix}' \mathbf{W}^{-1} \begin{pmatrix} \underline{y}_1 - \mathbf{H}_1(\cdot) \underline{y}_1 \\ \underline{y}_2 - \mathbf{H}_2(\cdot) \underline{y}_2 \end{pmatrix} \right] \\ &= (2\pi)^{-n} |\mathbf{W}|^{-1/2} \exp \left[-\frac{1}{2} \begin{pmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1 \\ (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2 \end{pmatrix}' \mathbf{W}^{-1} \begin{pmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1 \\ (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2 \end{pmatrix} \right]. \end{aligned} \quad (4.30)$$

Matriks kovarians *error* \mathbf{W} dalam persamaan (4.30) merupakan matriks bobot dengan struktur matriks seperti dinyatakan dalam persamaan (4.9), dan dapat dinyatakan dalam bentuk:

$$\mathbf{W} = \mathbf{\Sigma} \otimes \mathbf{I}, \quad (4.31)$$

dimana :

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_{11} & \vdots & \sigma_{12} \\ \vdots & \ddots & \vdots \\ \sigma_{12} & \vdots & \sigma_{22} \end{pmatrix}. \quad (4.32)$$

Dengan mensubstitusi persamaan (4.31) ke dalam persamaan (4.30), maka diperoleh:

$$L(\mathbf{\Sigma} | \underline{\alpha}, \underline{y}) = (2\pi)^{-n} |\mathbf{\Sigma} \otimes \mathbf{I}|^{-1/2} \exp \left[-\frac{1}{2} \begin{pmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1 \\ (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2 \end{pmatrix}' (\mathbf{\Sigma} \otimes \mathbf{I})^{-1} \begin{pmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1 \\ (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2 \end{pmatrix} \right]. \quad (4.33)$$

Dengan mengambil transformasi \ln , maka fungsi *log likelihood* dalam persamaan (4.33), dapat dinyatakan dengan persamaan sebagai berikut:

$$\ln L(\mathbf{\Sigma} | \underline{\alpha}, \underline{y}) = -n \ln(2\pi) - \frac{1}{2} \ln |\mathbf{\Sigma} \otimes \mathbf{I}| - \frac{1}{2} \begin{pmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1 \\ (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2 \end{pmatrix}' (\mathbf{\Sigma} \otimes \mathbf{I})^{-1} \begin{pmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1 \\ (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2 \end{pmatrix}. \quad (4.34)$$

Menggunakan Teorema 2.8.1.(i) dalam Gupta *et al.* (2000) diperoleh:

$$|\Sigma \otimes \mathbf{I}| = |\Sigma|^n |\mathbf{I}|^2 = |\Sigma|^n. \quad (4.35)$$

Dengan mensubstitusi persamaan (4.35) ke dalam persamaan (4.34), diperoleh:

$$\begin{aligned} \ln L(\Sigma | \alpha, y) &= -n \ln(2\pi) - \frac{1}{2} \ln |\Sigma|^n - \frac{1}{2} \left(\begin{pmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) y_1 \\ (\mathbf{I} - \mathbf{H}_2(\cdot)) y_2 \end{pmatrix}' (\Sigma \otimes \mathbf{I})^{-1} \begin{pmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) y_1 \\ (\mathbf{I} - \mathbf{H}_2(\cdot)) y_2 \end{pmatrix} \right) \\ &= -n \ln(2\pi) - \frac{n}{2} \ln |\Sigma| - \frac{1}{2} \left(\begin{pmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) y_1 \\ (\mathbf{I} - \mathbf{H}_2(\cdot)) y_2 \end{pmatrix}' (\Sigma \otimes \mathbf{I})^{-1} \begin{pmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) y_1 \\ (\mathbf{I} - \mathbf{H}_2(\cdot)) y_2 \end{pmatrix} \right) \quad (4.36) \end{aligned}$$

Apabila dalam persamaan (4.36), vektor:

$$\begin{pmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) y_1 \\ (\mathbf{I} - \mathbf{H}_2(\cdot)) y_2 \end{pmatrix}$$

dinyatakan dalam bentuk $\text{vec}(\mathbf{A})$ (vektor berorde $2n \times 1$), dimana

$\mathbf{A} = \begin{bmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) y_1 & (\mathbf{I} - \mathbf{H}_2(\cdot)) y_2 \end{bmatrix}$ adalah matriks berorde $n \times 2$, maka dengan

menggunakan Teorema 2.8.1.(ii) dalam Gupta *et al.* (2000) dan Proposisi 2.8.3 dalam Broxson (2006), diperoleh:

$$\begin{aligned} \begin{pmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) y_1 \\ (\mathbf{I} - \mathbf{H}_2(\cdot)) y_2 \end{pmatrix}' (\Sigma \otimes \mathbf{I})^{-1} \begin{pmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) y_1 \\ (\mathbf{I} - \mathbf{H}_2(\cdot)) y_2 \end{pmatrix} &= (\text{vec}(\mathbf{A}))' (\Sigma \otimes \mathbf{I})^{-1} \text{vec}(\mathbf{A}) \\ &= (\text{vec}(\mathbf{A}))' (\Sigma^{-1} \otimes \mathbf{I}) \text{vec}(\mathbf{A}) \\ &= \text{tr}(\Sigma^{-1} \mathbf{A}' \mathbf{A}) \\ &= \text{tr} \left(\Sigma^{-1} \begin{bmatrix} ((\mathbf{I} - \mathbf{H}_1(\cdot)) y_1)' \\ ((\mathbf{I} - \mathbf{H}_2(\cdot)) y_2)' \end{bmatrix} \begin{bmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) y_1 & (\mathbf{I} - \mathbf{H}_2(\cdot)) y_2 \end{bmatrix} \right). \quad (4.37) \end{aligned}$$

Dengan mensubstitusi persamaan (4.37) ke dalam persamaan (4.36), diperoleh:

$$\begin{aligned} \ln L(\Sigma | \alpha, y) &= -n \ln(2\pi) - \frac{n}{2} \ln |\Sigma| - \\ &\quad \frac{1}{2} \text{tr} \left(\Sigma^{-1} \begin{bmatrix} ((\mathbf{I} - \mathbf{H}_1(\cdot)) y_1)' \\ ((\mathbf{I} - \mathbf{H}_2(\cdot)) y_2)' \end{bmatrix} \begin{bmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) y_1 & (\mathbf{I} - \mathbf{H}_2(\cdot)) y_2 \end{bmatrix} \right). \quad (4.38) \end{aligned}$$

Nilai maksimum dari fungsi *likelihood* pada persamaan (4.38) diperoleh dengan menurunkan secara parsial $\ln L(\Sigma|\alpha, y)$ terhadap σ_{ij} ($i, j = 1, 2$), kemudian hasilnya disamakan dengan nol, sehingga diperoleh:

$$\begin{aligned} \frac{\partial}{\partial \sigma_{ij}} \ln L(\hat{\Sigma}|\alpha, y) &= -\frac{n}{2} \left(\frac{\partial}{\partial \sigma_{ij}} \ln |\hat{\Sigma}| \right) - \\ &\quad \frac{1}{2} \frac{\partial}{\partial \sigma_{ij}} \text{tr} \left(\hat{\Sigma}^{-1} \begin{bmatrix} ((\mathbf{I} - \mathbf{H}_1(\cdot))y_1)' \\ ((\mathbf{I} - \mathbf{H}_2(\cdot))y_2)' \end{bmatrix} \begin{bmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot))y_1 & (\mathbf{I} - \mathbf{H}_2(\cdot))y_2 \end{bmatrix} \right) = 0. \end{aligned} \quad (4.39)$$

Untuk mendapatkan turunan parsial dari persamaan (4.39), maka dapat digunakan proposisi 2.8.4 dalam Christensen (1991), sehingga diperoleh:

$$\frac{\partial}{\partial \sigma_{ij}} \text{tr} [\hat{\Sigma}^{-1}] = \text{tr} \left[\frac{\partial \hat{\Sigma}^{-1}}{\partial \sigma_{ij}} \right] \quad (4.40)$$

dan dengan menggunakan Proposisi 2.8.5 dalam Christensen (1991) :

$$\begin{aligned} \frac{\partial}{\partial \sigma_{ij}} \hat{\Sigma}^{-1} &= -\hat{\Sigma}^{-1} \frac{\partial \hat{\Sigma}}{\partial \sigma_{ij}} \hat{\Sigma}^{-1} \\ &= -\hat{\Sigma}^{-1} \mathbf{T}_{(i,j)} \hat{\Sigma}^{-1}, \end{aligned} \quad (4.41)$$

$$\begin{aligned} \frac{\partial}{\partial \sigma_{ij}} \ln |\hat{\Sigma}| &= \text{tr} \left[\hat{\Sigma}^{-1} \frac{\partial \hat{\Sigma}}{\partial \sigma_{ij}} \right] \\ &= \text{tr} [\hat{\Sigma}^{-1} \mathbf{T}_{(i,j)}], \end{aligned} \quad (4.42)$$

dimana $\mathbf{T}_{(i,j)}$ adalah matriks simetris berorde 2×2 yang elemennya bernilai 1 pada baris i kolom j dan baris j kolom i dan bernilai 0 untuk elemen lainnya.

Dengan menggunakan persamaan (4.40), (4.41), dan (4.42) ke dalam persamaan (4.39), maka diperoleh:

$$\begin{aligned} \frac{\partial}{\partial \sigma_{ij}} \ln L(\hat{\Sigma}|\alpha, y) &= -\frac{n}{2} \text{tr} [\hat{\Sigma}^{-1} \mathbf{T}_{(i,j)}] - \\ &\quad \frac{1}{2} \text{tr} \left(\frac{\partial}{\partial \sigma_{ij}} \left[\hat{\Sigma}^{-1} \begin{bmatrix} ((\mathbf{I} - \mathbf{H}_1(\cdot))y_1)' \\ ((\mathbf{I} - \mathbf{H}_2(\cdot))y_2)' \end{bmatrix} \begin{bmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot))y_1 & (\mathbf{I} - \mathbf{H}_2(\cdot))y_2 \end{bmatrix} \right] \right) = 0, \end{aligned}$$

$$\begin{aligned}
\frac{n}{2} \text{tr} \left[\hat{\Sigma}^{-1} \mathbf{T}_{(i,j)} \right] &= -\frac{1}{2} \text{tr} \left(\frac{\partial}{\partial \sigma_{ij}} \left[\hat{\Sigma}^{-1} \begin{bmatrix} ((\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1)' \\ ((\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2)' \end{bmatrix} \begin{bmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1 & (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2 \end{bmatrix} \right] \right) \\
\text{tr} \left(\hat{\Sigma}^{-1} \mathbf{T}_{(i,j)} \right) &= -\frac{1}{n} \text{tr} \left(\left(\frac{\partial \hat{\Sigma}^{-1}}{\partial \sigma_{ij}} \right) \begin{bmatrix} ((\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1)' \\ ((\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2)' \end{bmatrix} \begin{bmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1 & (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2 \end{bmatrix} \right) \\
&= -\frac{1}{n} \text{tr} \left(-\hat{\Sigma}^{-1} \mathbf{T}_{(i,j)} \hat{\Sigma}^{-1} \begin{bmatrix} ((\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1)' \\ ((\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2)' \end{bmatrix} \begin{bmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1 & (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2 \end{bmatrix} \right). \quad (4.43)
\end{aligned}$$

Karena n adalah skalar, maka persamaan (4.43) dapat dinyatakan dengan persamaan sebagai berikut:

$$\text{tr} \left(\hat{\Sigma}^{-1} \mathbf{T}_{(i,j)} \right) = \text{tr} \left(\hat{\Sigma}^{-1} \mathbf{T}_{(i,j)} \hat{\Sigma}^{-1} \frac{1}{n} \begin{bmatrix} ((\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1)' \\ ((\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2)' \end{bmatrix} \begin{bmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1 & (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2 \end{bmatrix} \right). \quad (4.44)$$

Kedua ruas dalam persamaan (4.44) akan bernilai sama dengan menetapkan:

$$\hat{\Sigma} = \frac{1}{n} \begin{bmatrix} ((\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1)' \\ ((\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2)' \end{bmatrix} \begin{bmatrix} (\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1 & (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2 \end{bmatrix}. \quad (4.45)$$

Hal ini dibuktikan dengan mensubstitusikan persamaan (4.45) ke dalam persamaan (4.44) sebagai berikut:

$$\begin{aligned}
\text{tr} \left(\hat{\Sigma}^{-1} \mathbf{T}_{(i,j)} \right) &= \text{tr} \left(\hat{\Sigma}^{-1} \mathbf{T}_{(i,j)} \hat{\Sigma}^{-1} \hat{\Sigma} \right) \\
&= \text{tr} \left(\hat{\Sigma}^{-1} \mathbf{T}_{(i,j)} \right).
\end{aligned}$$

Karena n adalah skalar, maka estimasi Σ pada persamaan (4.45) dapat dinyatakan dalam persamaan sebagai berikut:

$$\hat{\Sigma} = \begin{pmatrix} \frac{[(\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1]' [(\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1]}{n} & \frac{[(\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1]' [(\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2]}{n} \\ \frac{[(\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2]' [(\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1]}{n} & \frac{[(\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2]' [(\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2]}{n} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\underline{y}_1' (\mathbf{I} - \mathbf{H}_1(\cdot))' (\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1}{n} & \frac{\underline{y}_1' (\mathbf{I} - \mathbf{H}_1(\cdot))' (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2}{n} \\ \frac{\underline{y}_2' (\mathbf{I} - \mathbf{H}_2(\cdot))' (\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1}{n} & \frac{\underline{y}_2' (\mathbf{I} - \mathbf{H}_2(\cdot))' (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2}{n} \end{pmatrix}. \quad (4.46)$$

Karena $\underline{y}_1' (\mathbf{I} - \mathbf{H}_1(\cdot))' (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2$ adalah skalar maka nilainya sama dengan $\underline{y}_2' (\mathbf{I} - \mathbf{H}_2(\cdot))' (\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1$, dengan demikian maka estimasi $\hat{\Sigma}$ dalam persamaan (4.46) dapat dinyatakan kembali menjadi:

$$\hat{\Sigma} = \begin{pmatrix} \frac{\underline{y}_1' (\mathbf{I} - \mathbf{H}_1(\cdot))' (\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1}{n} & \frac{\underline{y}_1' (\mathbf{I} - \mathbf{H}_1(\cdot))' (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2}{n} \\ \frac{\underline{y}_1' (\mathbf{I} - \mathbf{H}_1(\cdot))' (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2}{n} & \frac{\underline{y}_2' (\mathbf{I} - \mathbf{H}_2(\cdot))' (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2}{n} \end{pmatrix}. \quad (4.47)$$

Berdasarkan persamaan (4.47), maka diperoleh estimasi matriks bobot \mathbf{W} dalam persamaan (4.31) sebagai berikut:

$$\begin{aligned} \hat{\mathbf{W}} &= \hat{\Sigma} \otimes \mathbf{I} \\ &= \begin{pmatrix} \hat{\sigma}_{11} & | & \hat{\sigma}_{12} \\ \hline \hat{\sigma}_{12} & | & \hat{\sigma}_{22} \end{pmatrix} \otimes \mathbf{I}, \end{aligned} \quad (4.48)$$

dengan:

$$\hat{\sigma}_{11} = \frac{\underline{y}_1' (\mathbf{I} - \mathbf{H}_1(\cdot))' (\mathbf{I} - \mathbf{H}_1(\cdot)) \underline{y}_1}{n},$$

$$\hat{\sigma}_{22} = \frac{\underline{y}_2' (\mathbf{I} - \mathbf{H}_2(\cdot))' (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2}{n},$$

$$\hat{\sigma}_{12} = \frac{\underline{y}_1' (\mathbf{I} - \mathbf{H}_1(\cdot))' (\mathbf{I} - \mathbf{H}_2(\cdot)) \underline{y}_2}{n}. \quad \blacksquare$$

Karena $\hat{\sigma}_{12} = \hat{\rho} \sqrt{\hat{\sigma}_{11} \hat{\sigma}_{22}}$, maka estimasi korelasi antar *error* ρ dapat dihitung dengan menggunakan persamaan (4.48) sebagai berikut:

$$\hat{\rho} = \frac{\hat{\sigma}_{12}}{\sqrt{\hat{\sigma}_{11} \hat{\sigma}_{22}}}$$

$$\begin{aligned}
& \frac{\underline{y}_1'(\mathbf{I} - \mathbf{H}_1(\cdot))'(\mathbf{I} - \mathbf{H}_2(\cdot))\underline{y}_2}{n} \\
= & \frac{\frac{\underline{y}_1'(\mathbf{I} - \mathbf{H}_1(\cdot))'(\mathbf{I} - \mathbf{H}_1(\cdot))\underline{y}_1}{n} \frac{\underline{y}_2'(\mathbf{I} - \mathbf{H}_2(\cdot))'(\mathbf{I} - \mathbf{H}_2(\cdot))\underline{y}_2}{n}}{\sqrt{\frac{\underline{y}_1'(\mathbf{I} - \mathbf{H}_1(\cdot))'(\mathbf{I} - \mathbf{H}_1(\cdot))\underline{y}_1}{n}} \sqrt{\frac{\underline{y}_2'(\mathbf{I} - \mathbf{H}_2(\cdot))'(\mathbf{I} - \mathbf{H}_2(\cdot))\underline{y}_2}{n}}} \\
= & \frac{\underline{y}_1'(\mathbf{I} - \mathbf{H}_1(\cdot))'(\mathbf{I} - \mathbf{H}_2(\cdot))\underline{y}_2}{\sqrt{\underline{y}_1'(\mathbf{I} - \mathbf{H}_1(\cdot))'(\mathbf{I} - \mathbf{H}_1(\cdot))\underline{y}_1} \sqrt{\underline{y}_2'(\mathbf{I} - \mathbf{H}_2(\cdot))'(\mathbf{I} - \mathbf{H}_2(\cdot))\underline{y}_2}}. \tag{4.49}
\end{aligned}$$

Halaman ini sengaja dikosongkan

BAB 5

HASIL KAJIAN SIMULASI

Bab ini merupakan kajian empiris yang membahas hasil tujuan penelitian 3 yang disebutkan dalam Bab 1, yaitu merancang studi simulasi pada model MARS birespon kontinu untuk melihat kebaikan dari model yang diperoleh. Hasil kajian empiris pada bab ini diperoleh berdasarkan tahapan metodologi dalam Bab 3 (Sub Bab 3.3), yang hasilnya diuraikan dalam subbab-subbab di bawah ini. Adapun studi simulasi yang akan dilakukan dalam penelitian ini meliputi studi simulasi MARS birespon kontinu pada fungsi asli MARS dan pada berbagai fungsi percobaan, seperti pada fungsi trigonometri, polinomial dan eksponensial.

5.1. Simulasi MARS Birespon Kontinu pada Fungsi Asli

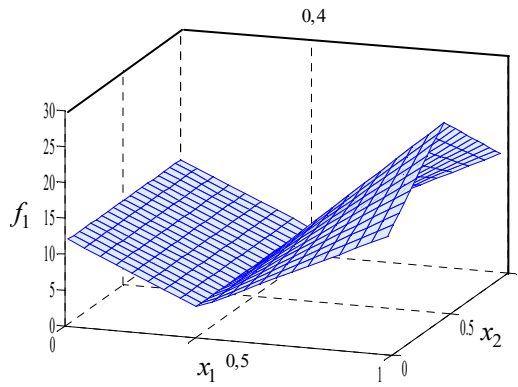
Sub bab berikut ini menguraikan hasil studi simulasi dalam mendapatkan estimasi model MARS birespon kontinu. Model yang disimulasikan dalam percobaan ini mengikuti bentuk asli dari fungsi regresi MARS, untuk mengetahui bagaimana kemampuan metode MARS birespon kontinu dalam mendapatkan hasil estimasi yang mendekati nilai fungsi aslinya. Contoh persamaan regresi yang dirancang untuk studi simulasi ini adalah sebagai berikut:

$$\begin{aligned} y_{1i} &= 4 + 39(x_{1i} - 0,4)_+ + 20(0,4 - x_{1i})_+ - 35(x_{1i} - 0,4)_+(x_{2i} - 0,5)_+ \\ &\quad - 37(x_{1i} - 0,4)_+(0,5 - x_{2i})_+ + \varepsilon_{1i}, \\ y_{2i} &= 5 + 37(x_{3i} - 0,45)_+ + 18(0,45 - x_{3i})_+ - 33(0,45 - x_{3i})_+(x_{4i} - 0,7)_+ \\ &\quad - 39(0,45 - x_{3i})_+(0,7 - x_{4i})_+ + \varepsilon_{2i}, \end{aligned} \quad (5.1)$$

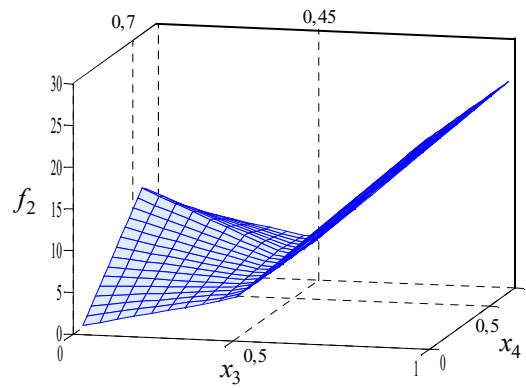
dimana prediktor x_{ji} ($j = 1, 2, 3, 4; i = 1, 2, \dots, 125$) dibangkitkan dari distribusi uniform: $x_{ji} \sim U(0,1)$, sedangkan *random error* $\varepsilon = (\varepsilon'_1, \varepsilon'_2)'$ dibangkitkan dari

distribusi normal: $\varepsilon \sim N(0, \mathbf{W})$. Matriks $\mathbf{W} = \begin{pmatrix} \sigma_{11}\mathbf{I} & \sigma_{12}\mathbf{I} \\ \sigma_{12}\mathbf{I} & \sigma_{22}\mathbf{I} \end{pmatrix}$, dengan varians *error*

$\sigma_{11} = \sigma_{22} = 1$, sedangkan kovarians *error* $\sigma_{12} = \rho\sqrt{\sigma_{11}\sigma_{22}}$ dengan korelasi antar *error* $\rho = 0,5$. Bentuk kurva regresi dari persamaan (5.1) disajikan dalam Gambar 5.1, masing-masing untuk respon 1 dan respon 2.



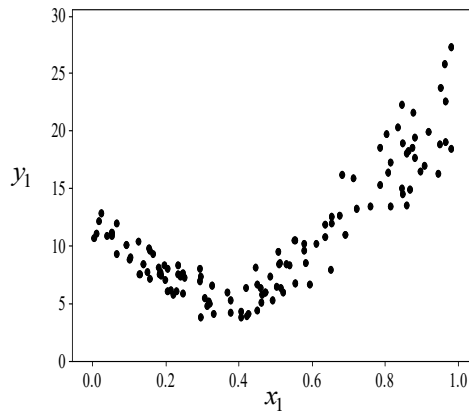
(a) Kurva regresi f_1



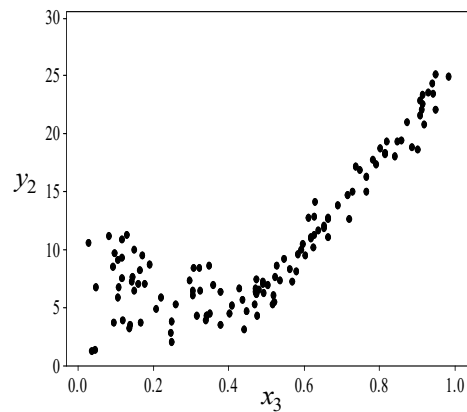
(b) Kurva regresi f_2

Gambar 5.1. Bentuk Kurva Regresi pada Simulasi Fungsi Asli MARS

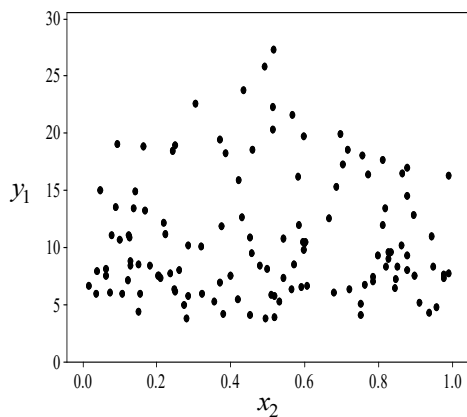
Hasil bangkitan data berdasarkan fungsi percobaan ini disajikan pada lampiran 10, sedangkan pola hubungan prediktor dengan respon disajikan pada Gambar 5.2.



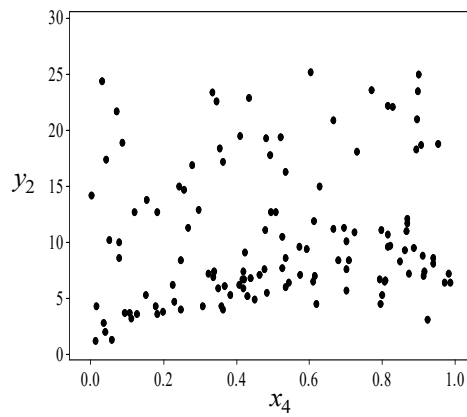
(a) *scatter plot* y_1 dengan x_1



(c) *scatter plot* y_2 dengan x_3



(b) *scatter plot* y_1 dengan x_2



(d) *scatter plot* y_2 dengan x_4

Gambar 5.2. *Scatter Plot* Respon dan Prediktor pada Simulasi Fungsi Asli MARS

Gambar 5.2 (a) dan (c) menunjukkan adanya perubahan pola hubungan prediktor dan respon dalam kurva regresi, selain itu Gambar 5.2 (b) dan (d) menunjukkan

pola hubungan yang tidak diketahui bentuknya. Hal ini mengindikasikan penggunaan regresi nonparametrik salah satunya MARS untuk estimasi model.

Estimasi fungsi regresi MARS birespon kontinu dalam simulasi ini diawali dengan mencari fungsi basis optimal dengan proses *stepwise* (*forward* dan *backward*) pada masing-masing respon. Hasil estimasi fungsi basis optimal untuk respon 1 dan respon 2 masing-masing disajikan dalam Tabel 5.1 dan Tabel 5.2.

Tabel 5.1. Estimasi Fungsi Basis Optimal pada Simulasi
Fungsi Asli MARS pada Respon 1

Fungsi Basis	Nilai Fungsi Basis
B_{10}	1
B_{11}	$(x_1 - 0,4069)_+$
B_{12}	$(0,4069 - x_1)_+$
B_{13}	$(x_1 - 0,4069)_+(x_2 - 0,5159)_+$
B_{14}	$(x_1 - 0,4069)_+(0,5159 - x_2)_+$

Tabel 5.2. Estimasi Fungsi Basis Optimal pada Simulasi
Fungsi Asli MARS pada Respon 2

Fungsi Basis	Nilai Fungsi Basis
B_{20}	1
B_{21}	$(x_3 - 0,4469)_+$
B_{22}	$(0,4469 - x_3)_+$
B_{23}	$(0,4469 - x_3)_+(x_4 - 0,7099)_+$
B_{24}	$(0,4469 - x_3)_+(0,7099 - x_4)_+$

Berdasarkan fungsi basis optimal dari masing-masing respon, maka diperoleh estimasi matriks bobot menggunakan persamaan (4.25) sebagai berikut:

$$\hat{\mathbf{W}} = \left(\begin{array}{cccc|cccc} 1,0865 & 0 & \cdots & 0 & 0,4637 & 0 & \cdots & 0 \\ 0 & 1,0865 & \cdots & 0 & 0 & 0,4637 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1,0865 & 0 & 0 & \cdots & 0,4637 \\ \hline 0,4637 & 0 & \cdots & 0 & 1,0989 & 0 & \cdots & 0 \\ 0 & 0,4637 & \cdots & 0 & 0 & 1,0989 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0,4637 & 0 & 0 & \cdots & 1,0989 \end{array} \right)$$

$$= \begin{pmatrix} 1,0865 & | & 0,4637 \\ \hline & & \\ 0,4637 & | & 1,0989 \end{pmatrix} \otimes \mathbf{I}, \quad (5.2)$$

dengan $\hat{\mathbf{W}}$ adalah estimasi matriks bobot berorde 250×250 dan \mathbf{I} adalah matriks identitas berorde 125. Berdasarkan estimasi matriks bobot dalam persamaan (5.2), maka diperoleh estimasi parameter model MARS birespon kontinu menggunakan persamaan (4.23), dengan melibatkan fungsi basis optimal dalam Tabel 5.1 dan Tabel 5.2 dengan hasil estimasi disajikan dalam Tabel 5.3.

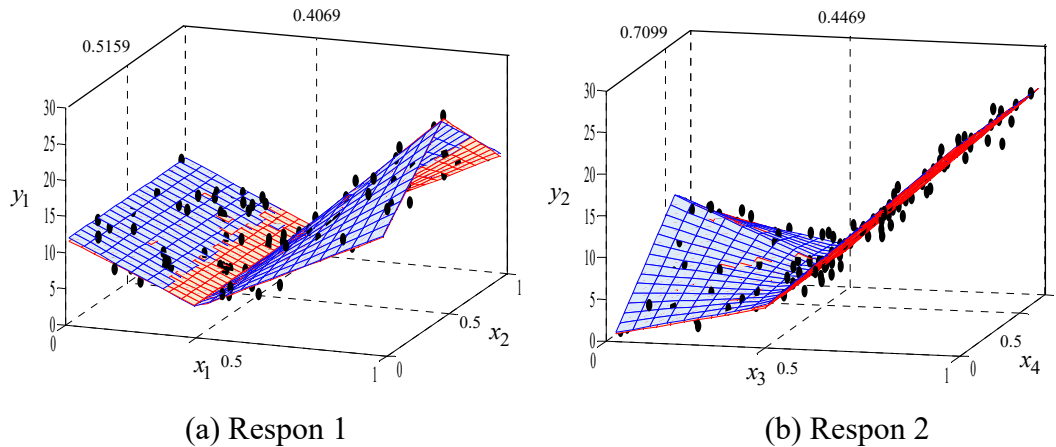
Tabel 5.3. Estimasi Parameter Model MARS Birespon Kontinu pada Simulasi Fungsi Asli MARS

Respon	Fungsi basis	Nilai Fungsi basis	Parameter
Respon 1	B ₁₀	1	4,04
	B ₁₁	$(x_{1i} - 0,4069)_+$	39,72
	B ₁₂	$(0,4069 - x_{1i})_+$	18,45
	B ₁₃	$(x_{1i} - 0,4069)_+ (x_{2i} - 0,5159)_+$	-38,72
	B ₁₄	$(x_{1i} - 0,4069)_+ (0,5159 - x_{2i})_+$	-37,46
Respon 2	B ₂₀	1	4,69
	B ₂₁	$(x_{3i} - 0,4469)_+$	37,29
	B ₂₂	$(0,4469 - x_{3i})_+$	18,01
	B ₂₃	$(0,4469 - x_{3i})_+ (x_{4i} - 0,7099)_+$	-28,98
	B ₂₄	$(0,4469 - x_{3i})_+ (0,7099 - x_{4i})_+$	-38,03

Berdasarkan estimasi parameter dan fungsi basis optimal dalam Tabel 5.3, maka diperoleh estimasi fungsi regresi MARS birespon kontinu dengan persamaan sebagai berikut:

$$\begin{aligned} \hat{f}_1(x_{1i}, x_{2i}) &= 4,04 + 39,72(x_{1i} - 0,4069)_+ + 18,45(0,4069 - x_{1i})_+ - (38,72(x_{1i} - 0,4069)_+ \\ &\quad \times (x_{2i} - 0,5159)_+ - 37,46(x_{1i} - 0,4069)_+ (0,5159 - x_{2i})_+), \\ \hat{f}_2(x_{3i}, x_{4i}) &= 4,69 + 37,29(x_{3i} - 0,4469)_+ + 18,01(0,4469 - x_{3i})_+ - (28,98(0,4469 - x_{3i})_+ \\ &\quad \times (x_{4i} - 0,7099)_+ - 38,03(0,4469 - x_{3i})_+ (0,7099 - x_{4i})_+). \end{aligned} \quad (5.3)$$

Persamaan fungsi regresi dalam persamaan (5.3) menunjukkan bahwa nilai estimasi parameter dan fungsi basis optimal mendekati nilai dari model asli atau model yang disimulasikan dalam persamaan (5.1). Nilai $RMSE$ dan R^2 yang dihasilkan oleh estimasi model MARS birespon kontinu dalam persamaan (5.3) masing-masing sebesar 1,0487 dan 96,62%. Bentuk grafik estimasi fungsi regresi disajikan dalam Gambar 5.3 masing-masing untuk respon 1 dan respon 2.

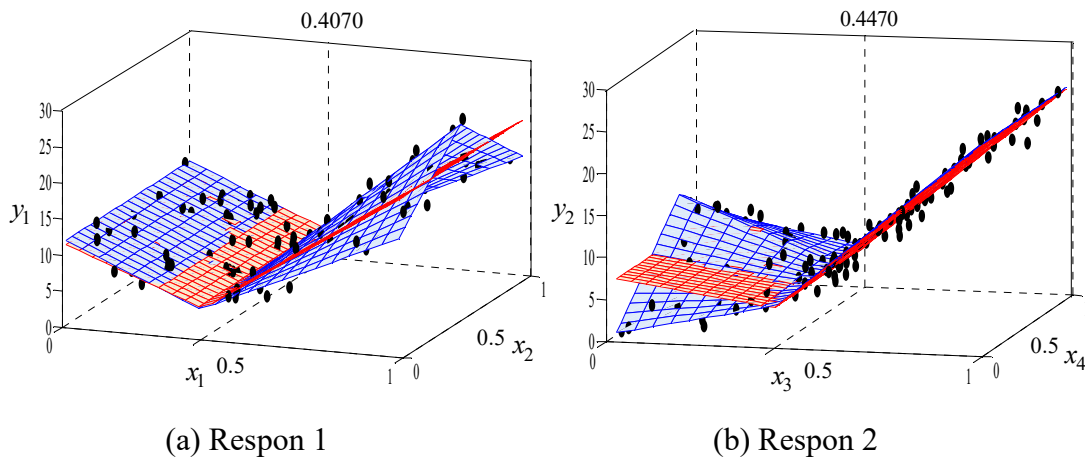


keterangan gambar: • observasi, fungsi asli dan estimasi fungsi regresi

Gambar 5.3. Grafik Estimasi Fungsi Regresi pada Simulasi
Fungsi Asli MARS dengan Fungsi Basis Optimal

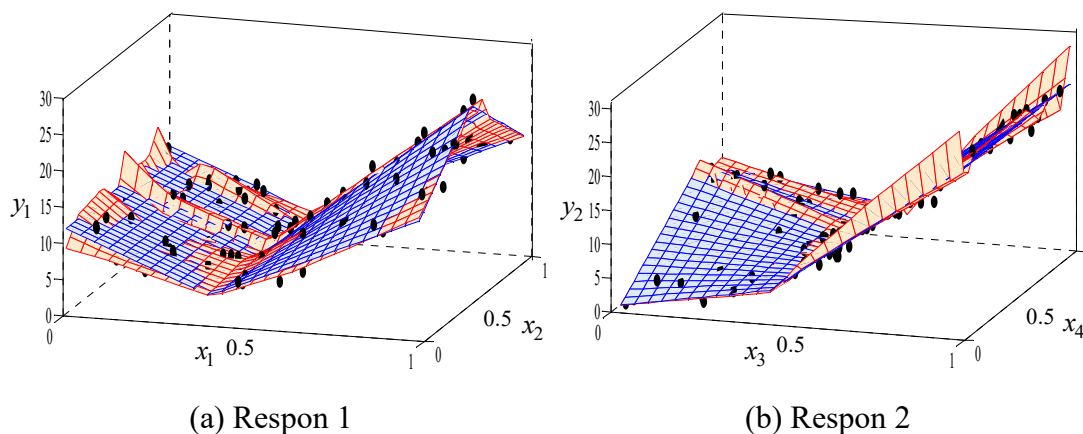
Fungsi basis optimal MARS birespon kontinu dalam persamaan (5.3) diambil dari Tabel 5.1 untuk respon 1 dan Tabel 5.2 untuk respon 2, yang dipilih dengan proses *stepwise* (*forward* dan *backward stepwise*) berdasarkan nilai GCV minimum pada masing-masing respon secara parsial. Dalam pemilihan fungsi basis optimal ini, diperoleh nilai GCV minimum untuk respon 1 sebesar 1,3063 dan untuk respon 2 sebesar 1,3213. Untuk mengakomodir korelasi antar respon, maka parameter regresi MARS birespon kontinu dalam persamaan (5.3) diperoleh dari hasil estimasi yang melibatkan matriks bobot dalam persamaan (5.2) dengan fungsi basis optimal pada masing-masing respon, dengan menggunakan persamaan (4.23). Sebagai perbandingan antara estimasi fungsi regresi yang memuat fungsi basis optimal dengan fungsi basis tidak optimal, maka disajikan bentuk estimasi fungsi regresi dengan sedikit fungsi basis dan banyak fungsi basis masing-masing pada Gambar 5.4 dan Gambar 5.5.

Gambar 5.4 menunjukkan bentuk estimasi fungsi regresi dengan sedikit fungsi basis (tidak optimal). Banyaknya fungsi basis untuk respon 1 adalah 3 fungsi basis dengan nilai GCV sebesar 3,3947 dan untuk respon 2 sebanyak 3 fungsi basis dengan nilai GCV sebesar 3,6252. Nilai $RMSE$ dan R^2 dari estimasi model MARS birespon ini masing-masing sebesar 1,7836 dan 90,22%. Dibandingkan dengan nilai $RMSE$ dan R^2 dari model optimal, tampak bahwa model ini memiliki keakuratan yang lebih rendah.



keterangan gambar: • observasi, ■ fungsi asli dan ■ estimasi fungsi regresi

Gambar 5.4. Grafik Estimasi Fungsi Regresi pada Simulasi
Fungsi Asli MARS dengan Sedikit Fungsi Basis



keterangan gambar: • observasi, ■ fungsi asli dan ■ estimasi fungsi regresi

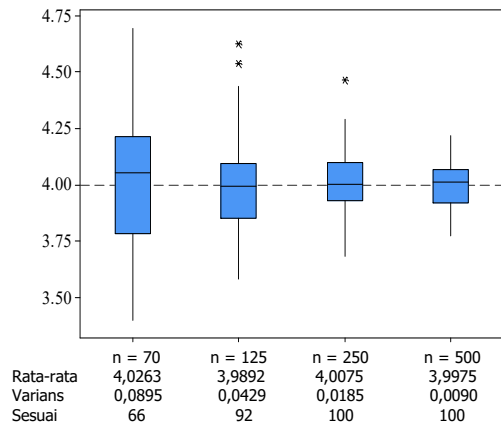
Gambar 5.5. Grafik Estimasi Fungsi Regresi pada Simulasi
Fungsi Asli MARS dengan Banyak Fungsi Basis

Gambar 5.5 menunjukkan bentuk estimasi fungsi regresi dengan banyak fungsi basis (tidak optimal). Banyaknya fungsi basis untuk respon 1 adalah 35 fungsi basis dengan nilai GCV sebesar 6,1363 dan untuk respon 2 sebanyak 35 fungsi basis dengan nilai GCV sebesar 6,5193. Nilai $RMSE$ dan R^2 dari estimasi model MARS birespon kontinu ini masing-masing sebesar 0,7868 dan 98,10%. Dibandingkan dengan nilai $RMSE$ dan R^2 dari model optimal, tampak bahwa model ini memiliki keakuratan yang lebih tinggi, penyebabnya karena fungsi basis yang termuat dalam model sangat banyak. Hal ini menyebabkan model yang dihasilkan sangat kompleks (tidak parsimoni). Model ini tidak baik untuk digunakan, karena memungkinkan munculnya kasus *overfitting* dalam regresi, yaitu baik untuk pemodelan data training namun tidak baik digunakan pada data *testing*.

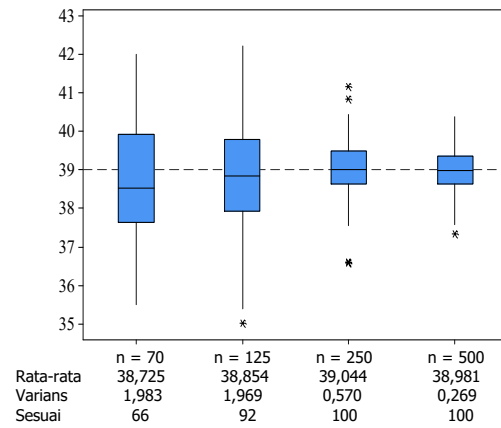
5.1.1. Studi Simulasi pada Variasi Ukuran Sampel dan Varians Error

Subbab berikut ini menunjukkan perilaku estimasi parameter (koefisien regresi) dan fungsi basis (knot) berdasarkan variasi ukuran sampel dan varians *error* pada model MARS birespon kontinu. Simulasi dilakukan dengan replikasi sebanyak 100 kali untuk melihat pola estimasi (parameter dan knot) yang dihasilkan oleh pengaruh variasi ukuran sampel dan varians *error* pada model tersebut. Studi simulasi dilakukan dengan membandingkan hasil estimasi parameter dan knot dengan nilai dari model aslinya (model yang disimulasikan), pada berbagai kondisi variasi ukuran sampel dan varians *error*. Hasil estimasi yang semakin mendekati nilai aslinya menunjukkan kemampuan estimasi model yang semakin baik.

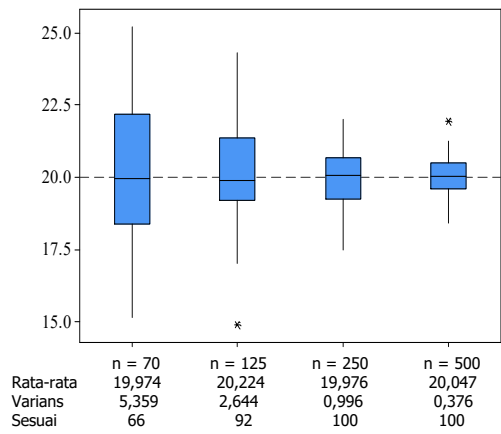
Gambar 5.6, 5.7, 5.8 dan 5.9 menyajikan contoh perbandingan estimasi parameter dan knot pada variasi ukuran sampel ($n = 70, 125, 250, 500$) masing-masing untuk respon 1 dan respon 2. Agar lebih mudah dalam membandingkan pengaruh ukuran sampel, maka dalam simulasi ini ditetapkan $\sigma_{11} = \sigma_{22}$. Pada pembahasan ini, diuraikan salah satu contoh simulasi untuk kondisi $\sigma_{11} = \sigma_{22} = 1$ dan korelasi antar *error* $\rho = 0,5$.



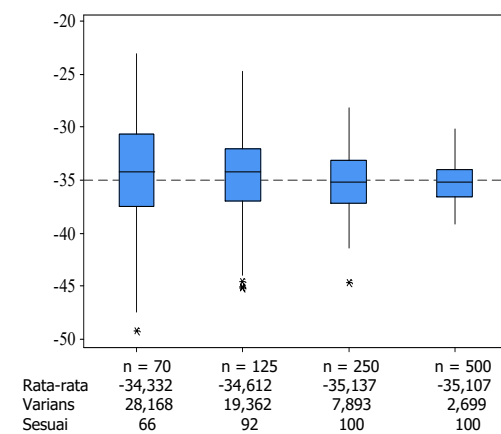
(a) Estimasi α_{10}



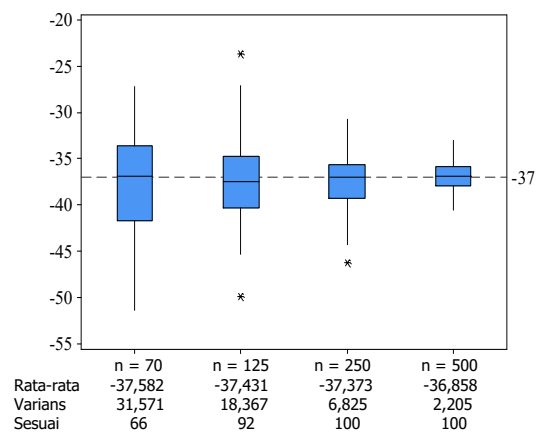
(b) Estimasi α_{11}



(c) Estimasi α_{12}

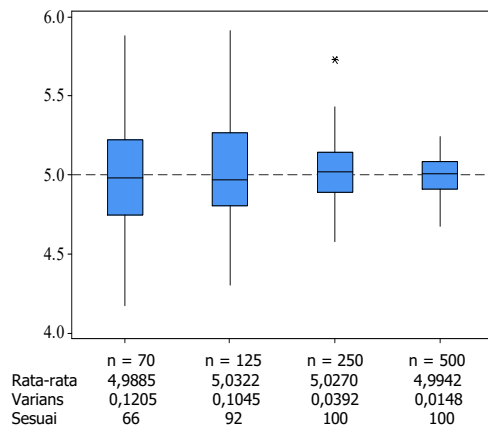


(d) Estimasi α_{13}

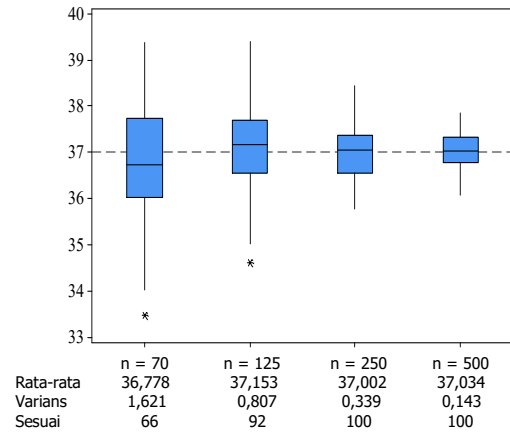


(e) Estimasi α_{14}

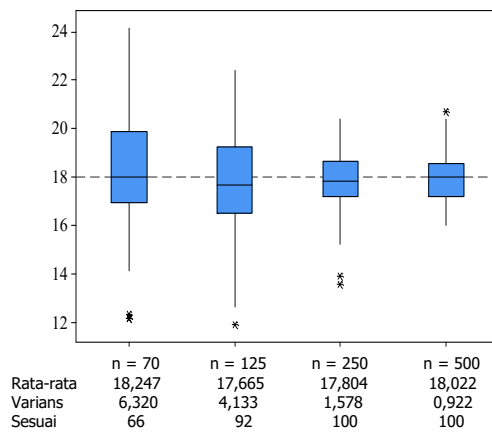
Gambar 5.6. *Boxplot* Estimasi Parameter untuk Berbagai Ukuran Sampel pada Simulasi Fungsi Asli MARS pada Respon 1



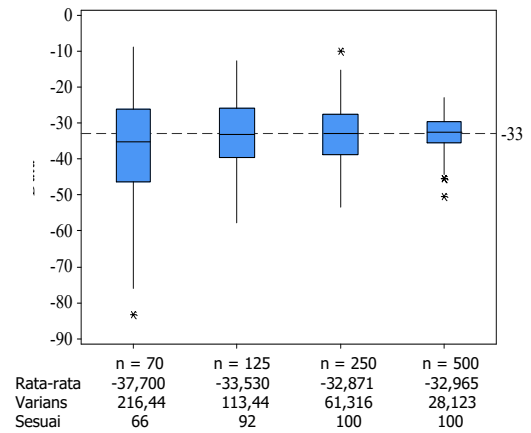
(a) Estimasi α_{20}



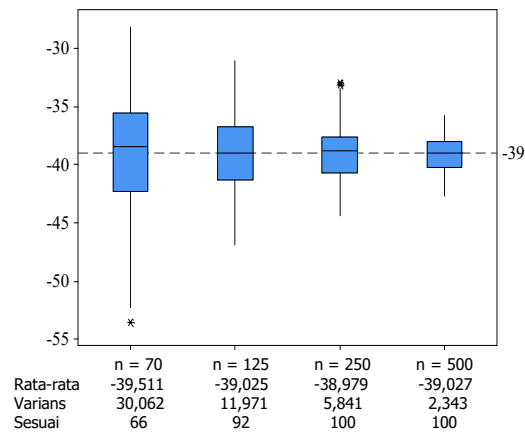
(b) Estimasi α_{21}



(c) Estimasi α_{22}

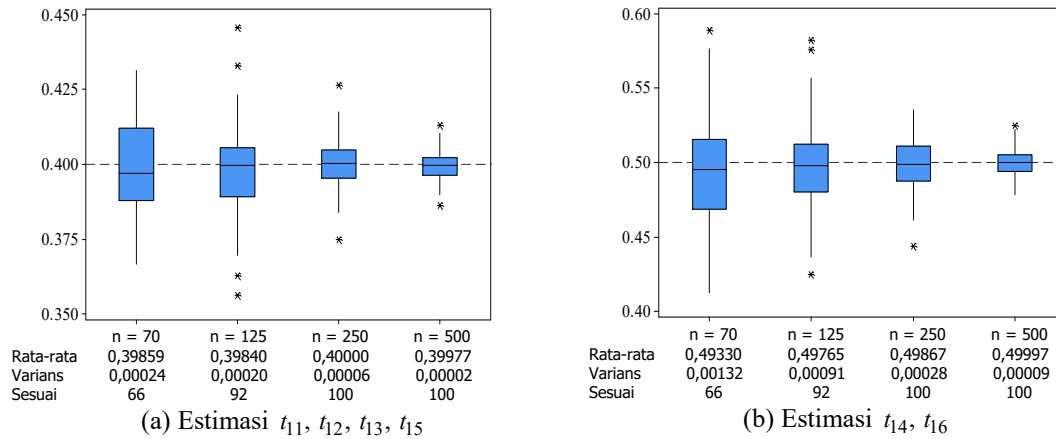


(d) Estimasi α_{23}

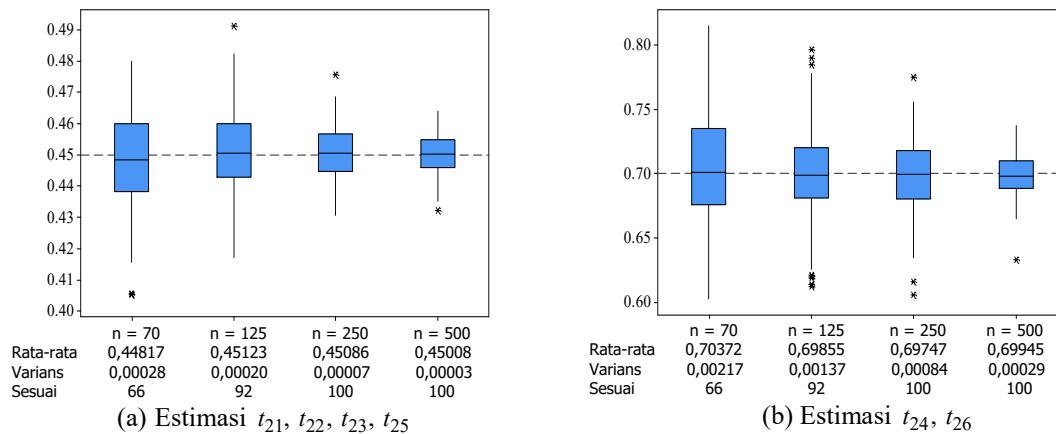


(e) Estimasi α_{24}

Gambar 5.7. *Boxplot* Estimasi Parameter untuk Berbagai Ukuran Sampel pada Simulasi Fungsi Asli MARS pada Respon 2



Gambar 5.8. *Boxplot* Estimasi Titik Knot untuk Berbagai Ukuran Sampel pada Simulasi Fungsi Asli MARS pada Respon 1

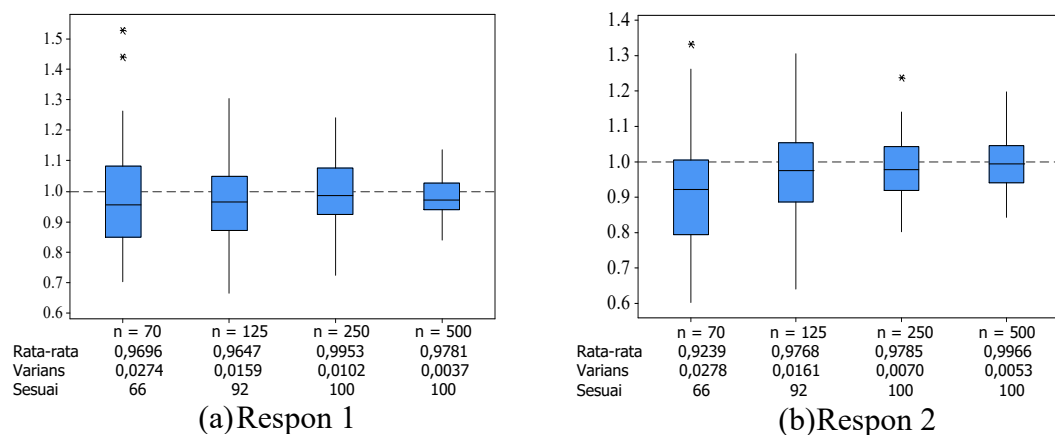


Gambar 5.9. *Boxplot* Estimasi Titik Knot untuk Berbagai Ukuran Sampel pada Simulasi Fungsi Asli MARS pada Respon 2

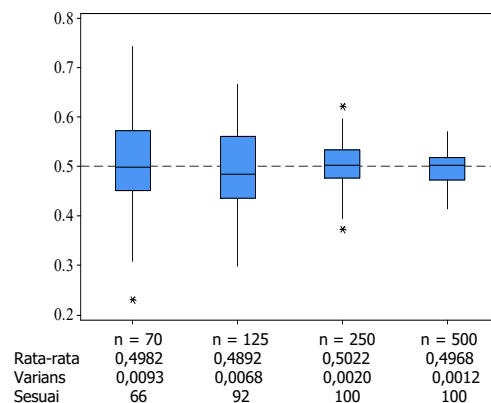
Hasil replikasi pada Gambar 5.6, 5.7, 5.8 dan 5.9 menunjukkan bahwa semakin besar ukuran sampel maka estimasi model yang diperoleh semakin mendekati model aslinya atau model yang disimulasikan. Hal ini ditunjukkan oleh semakin banyaknya hasil estimasi model (parameter dan knot) yang teridentifikasi sesuai dengan model aslinya, seiring dengan peningkatan ukuran sampel. Kondisi ini seperti ditunjukkan oleh Gambar 5.6, 5.7, 5.8 dan 5.9, bahwa pada $n = 70$, model yang teridentifikasi sesuai dengan model aslinya sebesar 66 model, kemudian meningkat pada $n = 125$ menjadi 92 model sesuai dan pada $n = 250$ dan $n = 500$ semua model teridentifikasi sesuai. Adapun estimasi model yang tidak teridentifikasi sesuai dengan model aslinya ($100 - \text{sesuai}$), disebabkan oleh

banyaknya fungsi basis hasil estimasi model yang tidak sama dengan banyaknya fungsi basis model aslinya. Hasil replikasi pada Gambar 5.6, 5.7, 5.8 dan 5.9 juga menunjukkan bahwa peningkatan ukuran sampel menghasilkan estimasi parameter dan knot yang umumnya lebih mendekati nilai aslinya dengan penyebaran (varians) hasil estimasi yang relatif lebih kecil (konsisten).

Variasi ukuran sampel juga berpengaruh pada hasil estimasi varians *error* dan korelasi *error* dari model yang diperoleh. Gambar 5.10 dan 5.11 menunjukkan bahwa estimasi varians *error* dan korelasi *error* dari model yang disimulasikan cenderung telah sesuai dengan varians *error* dan korelasi *error* dari model aslinya ($\sigma_{11} = \sigma_{22} = 1$ dan $\rho = 0,5$). Hasil replikasi pada Gambar 5.10 dan 5.11 menunjukkan bahwa peningkatan ukuran sampel menghasilkan estimasi varians *error* yang semakin mendekati nilai aslinya dengan penyebaran hasil estimasi yang relatif lebih kecil (memiliki rentang estimasi yang lebih kecil).

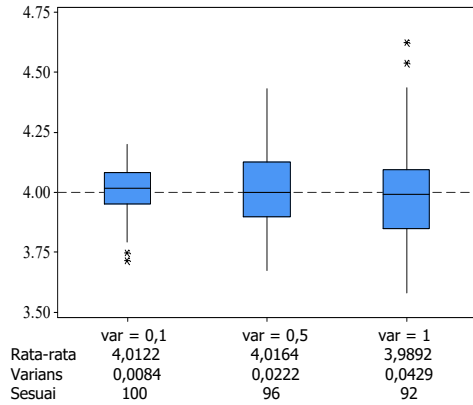


Gambar 5.10. *Boxplot* Estimasi Varians *Error* pada Berbagai Ukuran Sampel pada Simulasi Fungsi Asli MARS

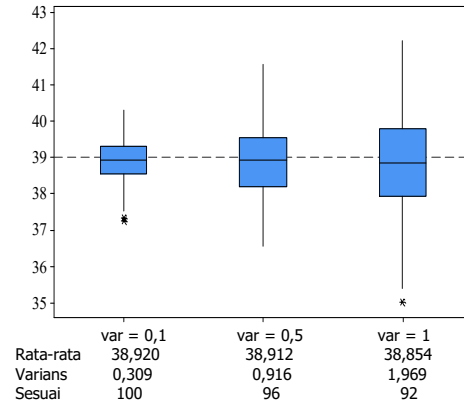


Gambar 5.11. *Boxplot* Estimasi Korelasi *Error* pada Berbagai Ukuran Sampel pada Simulasi Fungsi Asli MARS

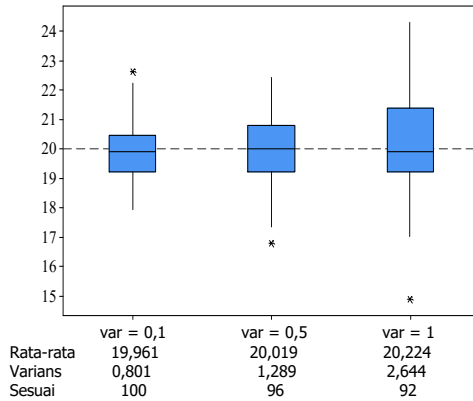
Selanjutnya, Gambar 5.12, 5.13, 5.14 dan 5.15 menyajikan contoh perbandingan estimasi parameter dan knot pada berbagai ukuran varians *error* ($\sigma_{11} = \sigma_{22} = 0,1$; $\sigma_{11} = \sigma_{22} = 0,5$; $\sigma_{11} = \sigma_{22} = 1$). Untuk memudahkan perbandingan estimasi, maka ditetapkan $n_1 = n_2$, dimana dalam simulasi ini $n_1 = n_2 = 125$.



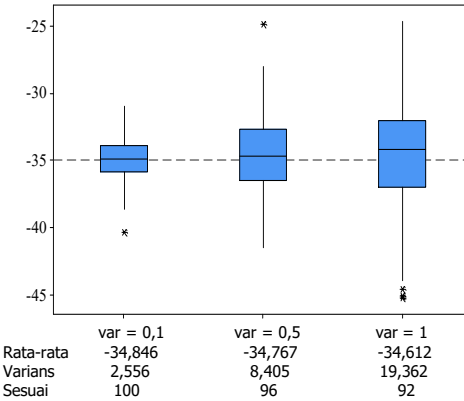
(a) Estimasi α_{10}



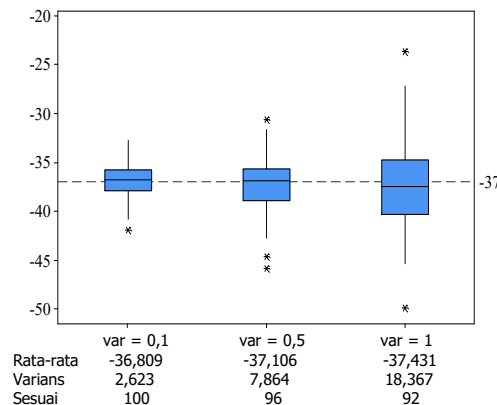
(b) Estimasi α_{11}



(c) Estimasi α_{12}

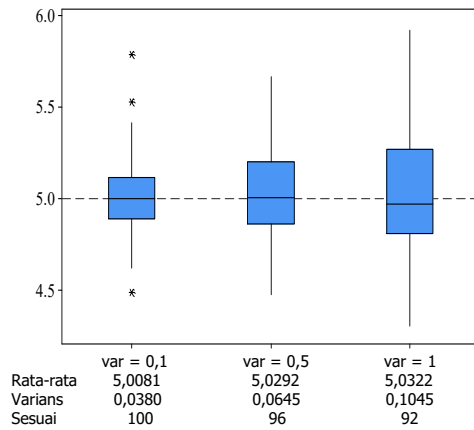


(d) Estimasi α_{13}

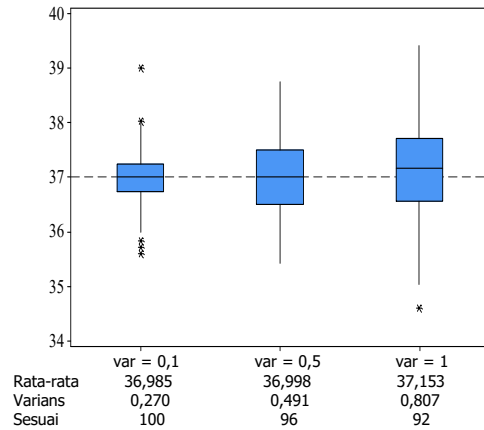


(e) Estimasi α_{14}

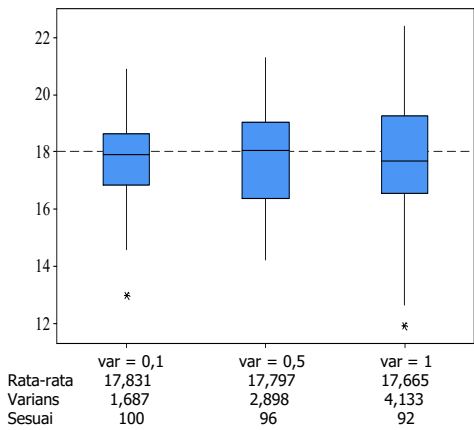
Gambar 5.12. *Boxplot* Estimasi Parameter untuk Berbagai Varians *Error* pada Simulasi Fungsi Asli MARS pada Respon 1



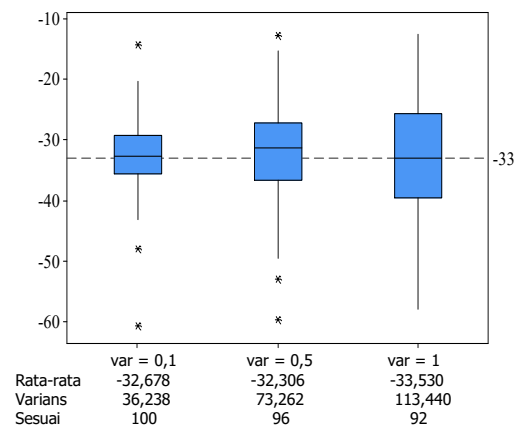
(a) Estimasi α_{20}



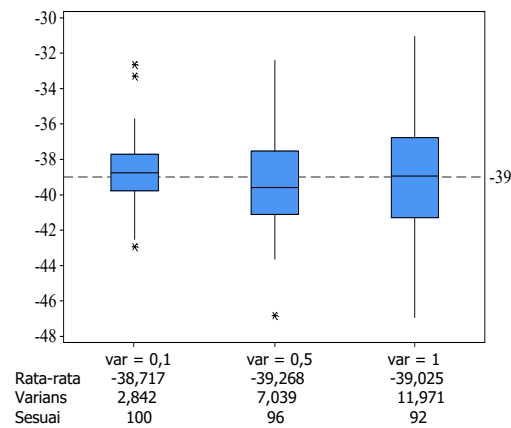
(b) Estimasi α_{21}



(c) Estimasi α_{22}

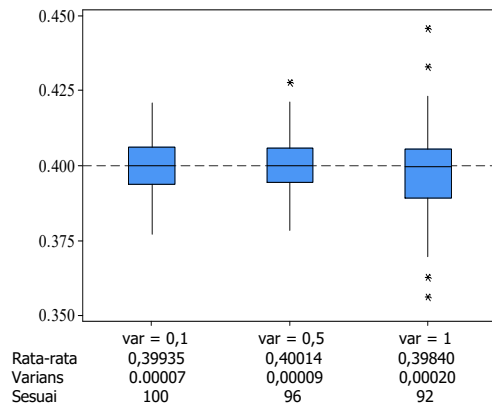


(d) Estimasi α_{23}

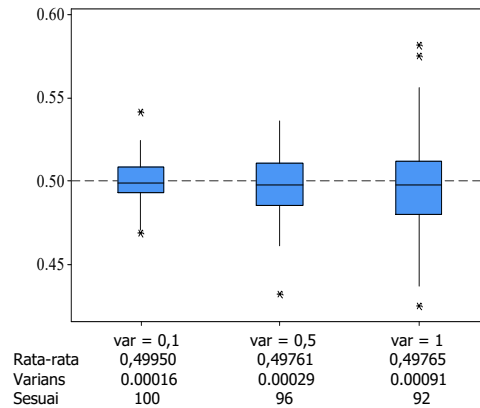


(e) Estimasi α_{24}

Gambar 5.13. *Boxplot* Estimasi Parameter untuk Berbagai Varians *Error* pada Simulasi Fungsi Asli MARS pada Respon 2

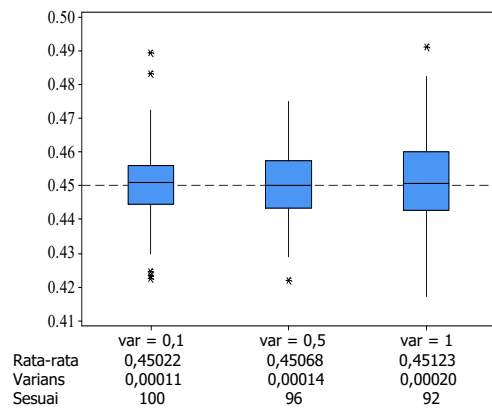


(a) Estimasi $t_{11}, t_{12}, t_{13}, t_{15}$

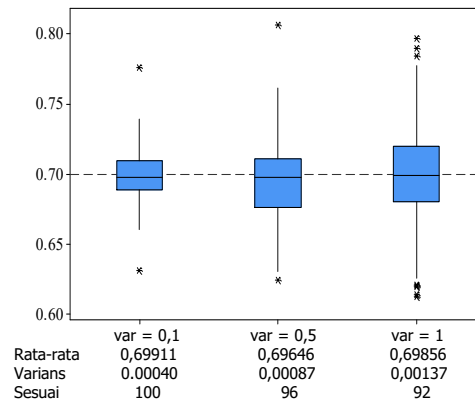


(b) Estimasi t_{14}, t_{16}

Gambar 5.14. *Boxplot* Estimasi Titik Knot untuk Berbagai Varians *Error* pada Simulasi Fungsi Asli MARS pada Respon 1



(a) Estimasi $t_{21}, t_{22}, t_{23}, t_{25}$



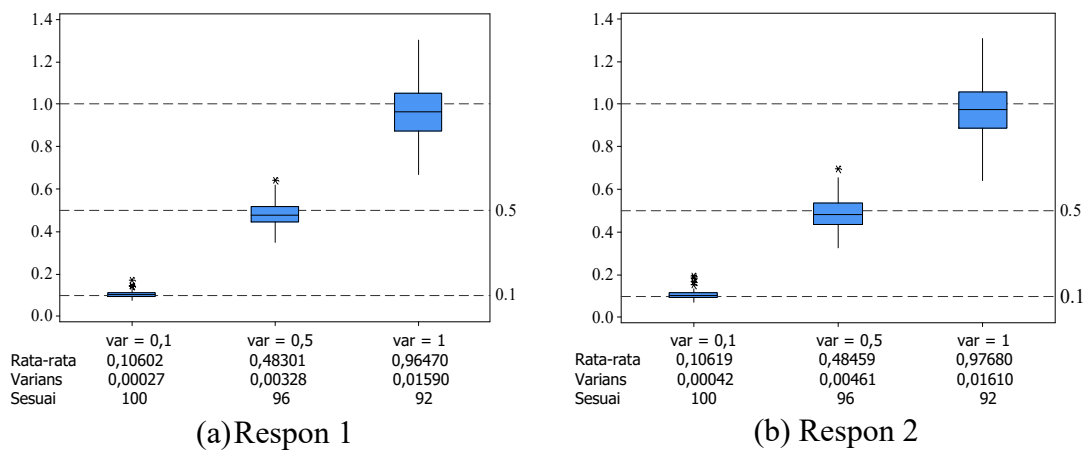
(b) Estimasi t_{24}, t_{26}

Gambar 5.15. *Boxplot* Estimasi Titik Knot untuk Berbagai Varians *Error* pada Simulasi Fungsi Asli MARS pada Respon 2

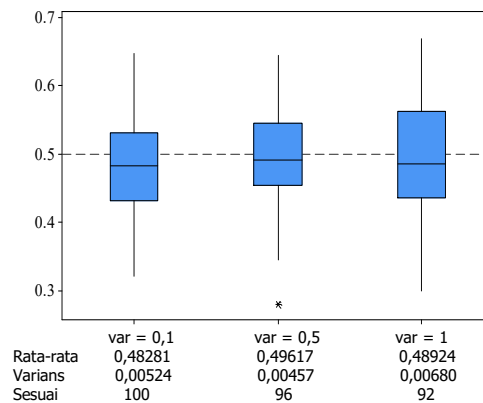
Hasil replikasi pada Gambar 5.12, 5.13, 5.14 dan 5.15 menunjukkan bahwa semakin kecil varians *error* sampel maka estimasi model yang diperoleh akan semakin mendekati model aslinya atau model yang disimulasikan. Hal ini ditunjukkan oleh semakin banyaknya hasil estimasi model (koefisien regresi dan knot) yang teridentifikasi sesuai dengan model yang sesungguhnya, seiring dengan penurunan varians *error*. Kondisi ini seperti ditunjukkan oleh Gambar 5.12, 5.13, 5.14 dan 5.15 bahwa pada varians = 0,1 model yang teridentifikasi sesuai dengan model aslinya sebesar 100 model, kemudian menurun pada varians = 0,5 menjadi 96 model sesuai dan pada varians = 1 menjadi 92 model teridentifikasi sesuai. Adapun estimasi model yang tidak teridentifikasi sesuai dengan model aslinya

(100 – sesuai), disebabkan oleh banyaknya fungsi basis hasil estimasi model yang tidak sama dengan banyaknya fungsi basis model aslinya. Hasil replikasi Gambar 5.12, 5.13, 5.14 dan 5.15 juga menunjukkan bahwa semakin kecil varians *error* data bangkitan menghasilkan estimasi parameter dan knot yang umumnya lebih mendekati nilai aslinya dengan penyebaran estimasi (varians) lebih kecil.

Gambar 5.16 dan 5.17 juga menunjukkan bahwa hasil estimasi varians *error* dan korelasi *error* dari model yang disimulasikan cenderung telah sesuai dengan nilai varians *error* dan korelasi *error* dari model aslinya. Hasil replikasi pada gambar tersebut menunjukkan bahwa semakin kecil ukuran varians *error* dari data bangkitan menghasilkan estimasi varians *error* model yang semakin mendekati nilai aslinya dengan penyebaran hasil estimasi (varians) yang relatif lebih kecil (memiliki rentang estimasi yang relatif lebih kecil).



Gambar 5.16. *Boxplot* Estimasi Varians *Error* untuk Berbagai Ukuran Varians *Error* pada Simulasi Fungsi Asli MARS



Gambar 5.17. *Boxplot* Estimasi Korelasi *Error* untuk Berbagai Ukuran Varians *Error* pada Simulasi Fungsi Asli MARS

Selanjutnya berdasarkan nilai rata-rata estimasi parameter dan knot dari 100 kali replikasi, seperti yang disajikan pada Gambar 5.12, 5.13, 5.14 dan 5.15, maka dapat dituliskan salah satu contoh fungsi regresi yang dihasilkan oleh ukuran sampel $n_1 = n_2 = 125$, varians error $\sigma_{11} = \sigma_{22} = 1$ dan korelasi antar error $\rho = 0,5$ sebagai berikut:

$$\begin{aligned}\hat{f}_1(x_{1i}, x_{2i}) &= 3,99 + 38,85(x_{1i} - 0,3984)_+ + 20,22(0,3984 - x_{1i})_+ - (34,61(x_{1i} - 0,3984)_+ \\ &\quad \times (x_{2i} - 0,4976)_+ - 37,43(x_{1i} - 0,3984)_+(0,4976 - x_{2i})_+, \\ \hat{f}_2(x_{3i}, x_{4i}) &= 5,03 + 37,15(x_{3i} - 0,4512)_+ + 17,66(0,4512 - x_{3i})_+ - (33,53(0,4512 - x_{3i})_+ \\ &\quad \times (x_{4i} - 0,6986)_+ - 39,02(0,4512 - x_{3i})_+(0,6986 - x_{4i})_+.\end{aligned}\tag{5.4}$$

Apabila dibandingkan dengan bentuk fungsi regresi dalam persamaan (5.1), tampak bahwa hasil estimasi fungsi regresi dalam persamaan (5.4) sudah mendekati bentuk aslinya. Hal ini menunjukkan bahwa estimator MARS birespon kontinu dapat bekerja dengan baik untuk mengestimasi bentuk fungsi regresi aslinya.

Untuk melihat pengaruh berbagai besaran ukuran sampel dan varians *error* (kecil, menengah, besar) pada model simulasi persamaan (5.1), maka dalam simulasi ini dilakukan perbandingan jumlah estimasi model yang sesuai. Hasil estimasi model dikatakan sesuai apabila jumlah fungsi basis hasil estimasi sama banyaknya dengan jumlah fungsi basis model aslinya, selain itu dikatakan sesuai apabila nilai estimasi parameter dan knot yang terdapat dalam estimasi model mendekati nilai parameter dan knot model aslinya. Simulasi dilakukan pada variasi ukuran sampel $n_\ell = 20, 50, 70, 125, 250, 500$ dan varians *error* $\sigma_{\ell\ell} = 0,1, 0,5, 1$ ($\ell = 1, 2$), dengan jumlah replikasi sebanyak 100 kali percobaan. Untuk memudahkan perbandingan estimasi model yang sesuai, maka ditetapkan $n_1 = n_2$ dan $\sigma_{11} = \sigma_{22}$.

Tabel 5.4 menunjukkan hasil simulasi 100 kali percobaan untuk melihat berapa banyak hasil estimasi model yang sesuai dengan model yang disimulasikan atau model aslinya. Tabel 5.4 menunjukkan bahwa semakin besar ukuran sampel (n) pada data simulasi, maka estimasi model yang diperoleh semakin banyak yang sesuai dengan model aslinya. Tabel 5.4 juga menunjukkan bahwa semakin

kecil nilai varians *error* pada data simulasi, maka estimasi model yang diperoleh semakin banyak yang sesuai dengan model aslinya.

Tabel 5.4. Jumlah Estimasi Model yang Sesuai dengan Model Aslinya pada Simulasi Fungsi Asli MARS dengan Replikasi 100 kali

Ukuran Sampel	Varians <i>Error</i> Data Bangkitan		
	0.1	0.5	1
20	6	0	0
30	35	12	3
40	62	38	14
50	82	59	36
60	95	72	54
70	99	78	66
125	100	96	92
250	100	100	100
500	100	100	100

Hasil replikasi pada Tabel 5.4 juga menunjukkan bahwa metode MARS birespon kontinu cenderung baik digunakan untuk ukuran sampel (n) diatas 50 observasi. Hal ini sejalan dengan Friedman (1991), yang menyarankan penggunaan metode MARS pada observasi $50 \leq n \leq 1000$. Sedangkan pada ukuran sampel $n < 50$ dengan varians error yang besar cenderung memberikan hasil yang kurang baik.

5.1.2. Studi Simulasi pada Variasi Ukuran Korelasi

Hal yang perlu diperhatikan dalam pemodelan data adalah bagaimana membentuk model yang sesuai/tepat menggambarkan fenomena riil yang kompleks bukan mendapatkan model yang sempurna. Salah satu cara untuk mengetahuinya adalah dengan melakukan validasi model untuk melihat ketepatan model dalam prediksi. Ketepatan prediksi dalam MARS dapat dievaluasi dengan *Root Mean Square Error Prediction (RMSEP)* (Jekabsons, 2015 dan Milborrow, 2015). Dalam MARS, nilai *RMSEP* dihitung dengan cara membagi data menjadi data *training* dan data *testing* menggunakan *k-fold cross-validation*. *K-fold*

membagi data menjadi k -grup dengan jumlah yang sama. Setiap data grup (data *testing*) diprediksi menggunakan model yang dibangun dari data di luar grup (data *training*) dan dihitung *RMSEP* nya. Selanjutnya nilai *RMSEP* model diperoleh dari rata-rata *RMSEP* dari keseluruhan grup (k -grup).

Untuk mengetahui pengaruh korelasi dalam ketepatan prediksi model MARS birespon kontinu, maka dalam penelitian ini dibandingkan nilai *RMSEP* dari estimasi model (persamaan (4.17)) yang tidak melibatkan korelasi dalam matriks bobot:

$$\mathbf{W} = \begin{pmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{pmatrix} \otimes \mathbf{I}, \quad (5.5)$$

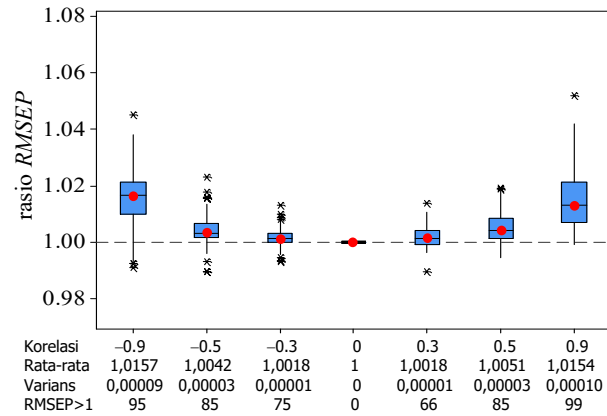
dengan *RMSEP* dari estimasi model MARS birespon kontinu (persamaan (4.17)) yang melibatkan korelasi dalam matriks bobot:

$$\mathbf{W} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \otimes \mathbf{I}, \quad (5.6)$$

dimana $\sigma_{12} = \rho \sqrt{\sigma_{11} \sigma_{22}}$. Contoh k -fold yang digunakan dalam simulasi ini sama dengan 10.

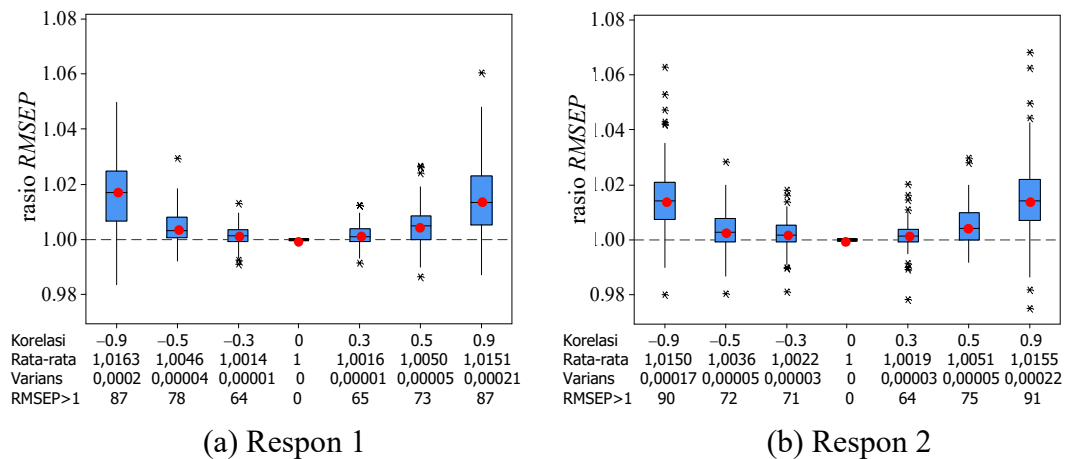
Untuk melihat pola nilai rasio *RMSEP* birespon yang dihasilkan oleh variasi besaran ukuran korelasi, maka simulasi dilakukan dengan replikasi sebanyak 100 kali. Gambar 5.18 menyajikan contoh perbandingan nilai rasio *RMSEP* birespon untuk berbagai ukuran korelasi ($\rho = -0,9, -0,5, -0,3, 0, 0,3, 0,5, 0,9$) pada model persamaan (5.1). Untuk memudahkan dalam membandingkan nilai rasio *RMSEP* birespon pada berbagai ukuran korelasi, maka dalam simulasi ini ditetapkan $n_1 = n_2$ dan $\sigma_{11} = \sigma_{22}$, dimana pada simulasi ini diuraikan salah satu contoh untuk $n_1 = n_2 = 125$ dan $\sigma_{11} = \sigma_{22} = 1$.

Gambar 5.18 menunjukkan bahwa semakin meningkat nilai mutlak korelasi ($|\rho|$), maka semakin besar penyebaran (varians) nilai rasio *RMSEP* birespon yang dihasilkan. Nilai rasio *RMSEP* birespon menurut gambar tersebut umumnya berada diatas nilai 1 seiring dengan meningkatnya nilai mutlak korelasi. Hal ini berarti bahwa nilai *RMSEP* birespon dari model yang melibatkan korelasi umumnya menghasilkan nilai *RMSEP* model yang lebih kecil daripada tidak melibatkan korelasi, seiring dengan meningkatnya nilai mutlak korelasi.



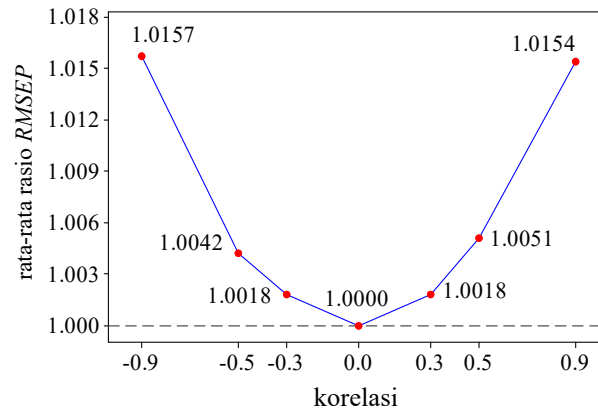
Gambar 5.18. *Boxplot* Rasio *RMSEP* Birespon berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Asli MARS

Apabila hasil pemodelan MARS birespon kontinu diuraikan untuk setiap respon, ternyata nilai rasio *RMSEP* untuk respon 1 dan respon 2 pada Gambar 5.19 menunjukkan pola yang sama dengan nilai rasio *RMSEP* birespon yang disajikan pada Gambar 5.18. Hal ini memperjelas bukti bahwa besarnya nilai mutlak korelasi memberikan pengaruh pada penurunan nilai *RMSEP* birespon yang dihasilkan.



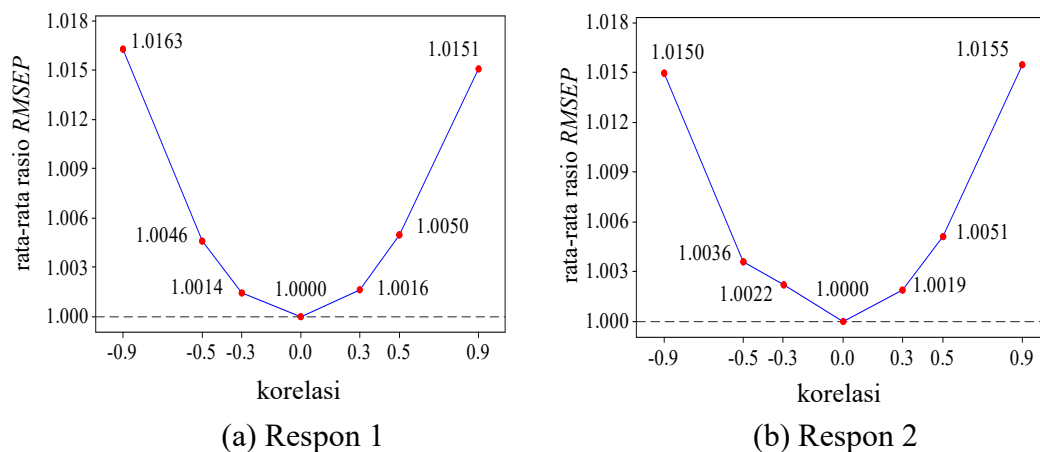
Gambar 5.19. *Boxplot* Rasio *RMSEP* berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Asli MARS pada Setiap Respon

Selanjutnya berdasarkan data *boxplot* pada Gambar 5.18, maka dapat disajikan perbandingan rata-rata nilai rasio *RMSEP* birespon pada berbagai ukuran korelasi dalam Gambar 5.20.



Gambar 5.20. Rata-rata Rasio $RMSEP$ Birespon berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Asli MARS

Perbandingan rata-rata nilai rasio $RMSEP$ birespon pada berbagai ukuran korelasi juga dapat diuraikan untuk setiap respon berdasarkan data *boxplot* Gambar 5.19, yang hasilnya disajikan pada Gambar 5.21.



Gambar 5.21. Rata-rata Rasio $RMSEP$ berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Asli MARS pada Setiap Respon

Gambar 5.20 dan 5.21 secara jelas menunjukkan pengaruh variasi besaran ukuran korelasi pada rata-rata nilai $RMSEP$ model MARS birespon kontinu. Hasil simulasi pada gambar tersebut menunjukkan bahwa semakin meningkat nilai mutlak korelasi pada model birespon, maka semakin besar rata-rata nilai rasio $RMSEP$ birespon yang dihasilkan. Hal ini menjelaskan bahwa semakin besar nilai mutlak korelasi dalam model MARS birespon kontinu, maka

semakin kecil nilai *RMSEP* dari model yang melibatkan korelasi dibandingkan dengan model yang tidak melibatkan korelasi, yang berarti bahwa hasil prediksinya semakin lebih baik. Berdasarkan rata-rata nilai rasio *RMSEP* juga menunjukkan bahwa nilai mutlak korelasi diatas 0,3 cenderung memberikan hasil prediksi MARS birespon kontinu yang lebih baik.

Berdasarkan uraian simulasi pada Fungsi asli, diperoleh beberapa kesimpulan sebagai berikut:

1. Semakin besar ukuran sampel maka semakin dekat nilai estimasi parameter dan knot dengan nilai aslinya dengan penyebaran/varians hasil estimasi yang relatif lebih kecil (konsisten), sebaliknya semakin kecil ukuran sampel maka semakin jauh nilai estimasi parameter dan knot dibandingkan dengan nilai aslinya.
2. Semakin kecil varians *error* maka semakin dekat nilai estimasi parameter dan knot dengan nilai aslinya dengan penyebaran hasil estimasi yang relatif lebih kecil, sebaliknya semakin besar varians *error* maka semakin jauh nilai estimasi parameter dan knot dibandingkan dengan nilai aslinya.
3. Metode MARS birespon kontinu cenderung baik digunakan untuk ukuran sampel (n) diatas 50 observasi ($n \geq 50$), dan cenderung kurang baik digunakan untuk $n < 50$ dengan varians *error* yang besar.
4. Jika fungsi basis yang digunakan sangat sedikit, maka tingkat keakuratan model akan sangat rendah, yang ditandai oleh tingginya nilai *RMSE* dan rendahnya nilai R^2 model. Sebaliknya jika fungsi basis yang digunakan sangat banyak, maka model yang dihasilkan akan sangat kompleks. Akibatnya perlu dipilih fungsi basis yang optimal yang paling sesuai untuk data.
5. Semakin besar nilai mutlak korelasi antar respon maka semakin baik penggunaan estimator MARS birespon, karena semakin rendah nilai *RMSEP* model yang diperoleh. Simulasi juga menunjukkan bahwa nilai mutlak korelasi diatas 0,3 ($|\rho| > 0,3$) cenderung memberikan hasil prediksi yang lebih baik.

5.2. Simulasi MARS Birespon Kontinu pada Berbagai Fungsi Percobaan

Sub bab ini menguraikan hasil simulasi MARS birespon pada berbagai fungsi percobaan. Simulasi dilakukan pada kombinasi fungsi Trigonometri-Polinomial, Eksponensial-Trigonometri dan Polinomial-Eksponensial.

5.2.1. Simulasi MARS Birespon pada Fungsi Trigonometri-Polinomial

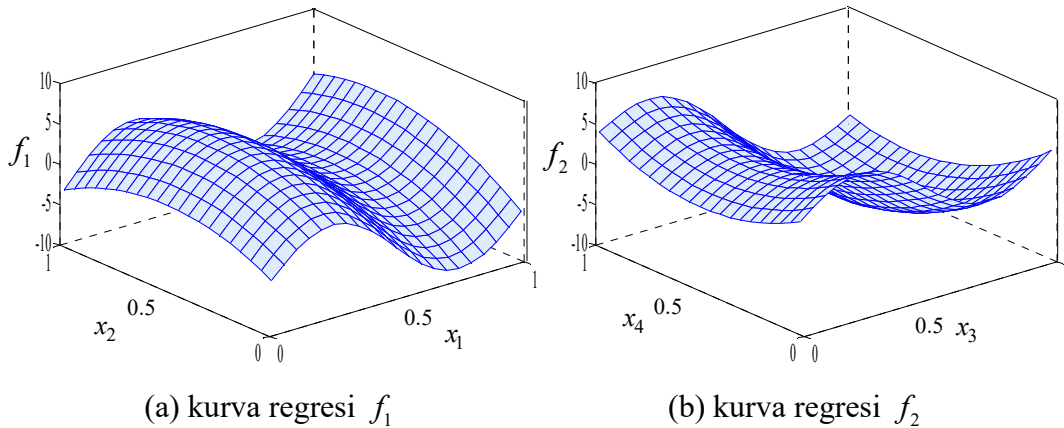
Bagian sub bab ini menguraikan hasil simulasi MARS birespon kontinu pada kombinasi fungsi trigonometri dan polinomial. Simulasi dilakukan untuk mengetahui kemampuan MARS birespon kontinu dalam mendapatkan hasil estimasi yang mendekati bentuk fungsi trigonometri-polinomial melalui pendekatan grafik. Contoh persamaan regresi yang dirancang untuk studi simulasi ini adalah sebagai berikut:

$$\begin{aligned} y_{1i} &= 5 \sin(2\pi x_{1i}) - 15(0,5 - x_{2i})^2 + 6x_{1i}x_{2i} + \varepsilon_{1i}, \\ y_{2i} &= 4 \sin(2\pi x_{3i}) + 16(x_{4i} - 0,5)^2 - 7x_{3i}x_{4i} + \varepsilon_{2i}, \end{aligned} \quad (5.7)$$

dimana prediktor x_{ji} ($j = 1, 2, 3, 4; i = 1, 2, \dots, 125$) dibangkitkan dari distribusi uniform: $x_{ji} \sim U(0,1)$, sedangkan *random error* $\varepsilon = (\varepsilon'_1, \varepsilon'_2)'$ dibangkitkan dari

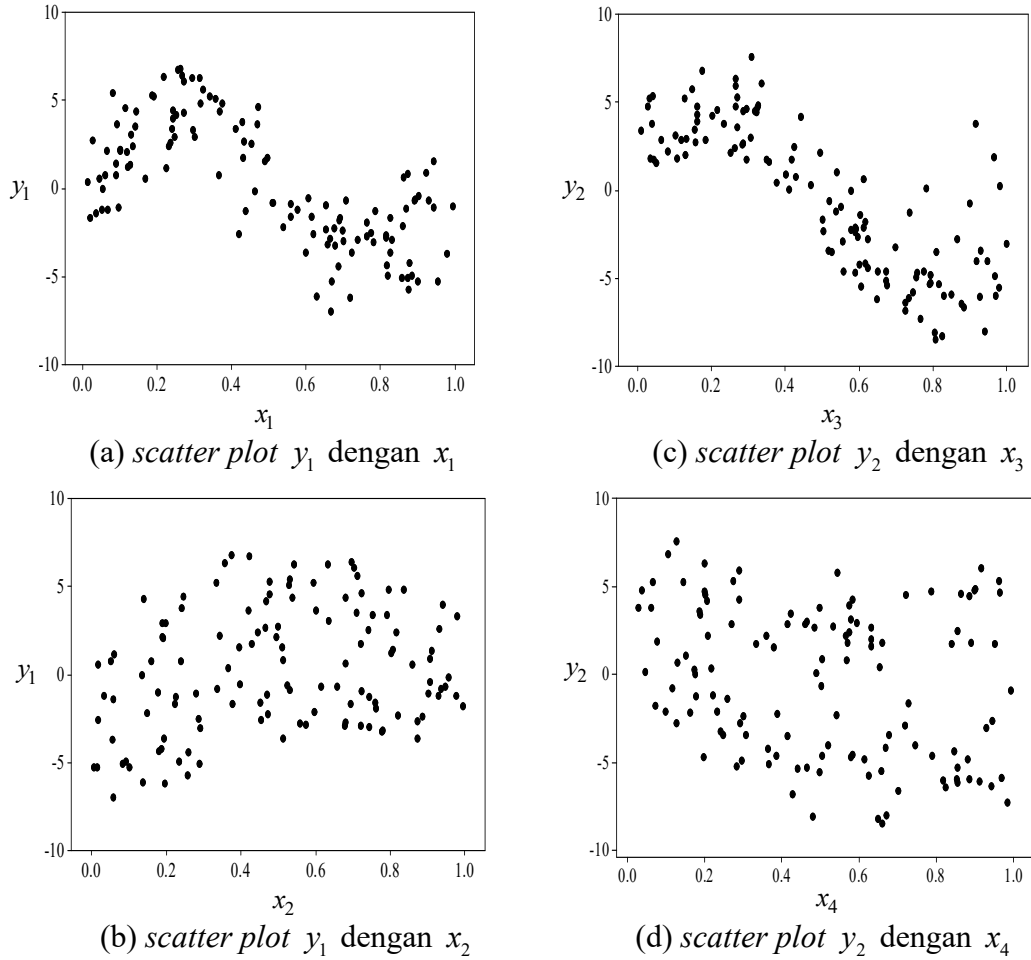
distribusi normal: $\varepsilon \sim N(0, \mathbf{W})$. Matriks $\mathbf{W} = \begin{pmatrix} \sigma_{11}\mathbf{I} & \sigma_{12}\mathbf{I} \\ \sigma_{12}\mathbf{I} & \sigma_{22}\mathbf{I} \end{pmatrix}$, dimana varians *error*

$\sigma_{11} = \sigma_{22} = 1$, sedangkan kovarians *error* $\sigma_{12} = \rho\sqrt{\sigma_{11}\sigma_{22}}$ dengan korelasi antar *error* $\rho = 0,5$. Bentuk kurva regresi dari persamaan (5.7) disajikan dalam Gambar 5.22, masing-masing untuk respon 1 dan respon 2.



Gambar 5.22. Bentuk Kurva Regresi pada Simulasi Fungsi Trigonometri-Polinomial

Hasil bangkitan prediktor dan respon berdasarkan fungsi percobaan dalam persamaan (5.7) ini disajikan pada data lampiran 11, dengan pola hubungan antara prediktor dengan respon disajikan dalam Gambar 5.23.



Gambar 5.23. Scatter Plot Respon dengan Prediktor pada Simulasi Fungsi Trigonometri-Polinomial

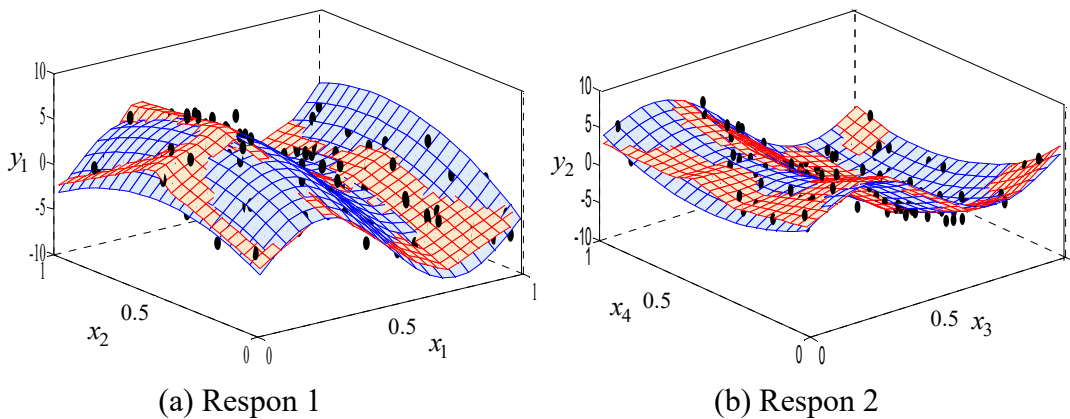
Gambar 5.23 menunjukkan pola hubungan antar respon dan prediktor yang tidak diketahui bentuknya. Hal ini mengindikasikan penggunaan regresi nonparametrik salah satunya MARS birespon kontinu untuk mengestimasi model. Hasil estimasi model (estimasi parameter dan fungsi basis optimal) disajikan dalam Tabel 5.5. Sehingga berdasarkan Tabel 5.5 diperoleh persamaan regresi sebagai berikut:

$$\begin{aligned}\hat{f}_1(x_{1i}, x_{2i}) &= 17,08 - 10,24(0,53 - x_{2i})_+ + 56,14(x_{1i} - 0,72)_+ - 28,46(0,72 - x_{1i})_+ - \\ &\quad 85,83(x_{2i} - 0,53)_+(0,32 - x_{1i})_+ - 42,43(x_{1i} - 0,27)_+ + (84,66(0,19 - x_{1i})_+ \\ &\quad \times (x_{2i} - 0,44)_+) - 2,18(x_{2i} - 0,37)_+ + 10,3(0,60 - x_{1i})_+, \\ \hat{f}_2(x_{3i}, x_{4i}) &= -9,94 + 24,97(0,83 - x_{3i})_+ + 13,95(0,38 - x_{4i})_+ + (15,75(0,83 - x_{3i})_+ \times \\ &\quad (x_{4i} - 0,66)_+) - 51,06(0,25 - x_{3i})_+(x_{4i} - 0,58)_+ - (33,64(0,25 - x_{3i})_+ \times \\ &\quad (0,58 - x_{4i})_+) - 21,58(0,32 - x_{3i})_+ + 41,79(x_{3i} - 0,65)_+(0,67 - x_{4i})_+ + \\ &\quad 23,69(x_{3i} - 0,65)_+(x_{4i} - 0,13)_+.\end{aligned}\tag{5.8}$$

Tabel 5.5. Estimasi Parameter dan Fungsi Basis Optimal MARS
Birespon pada Simulasi Fungsi Trigonometri-Polinomial

Respon	Fungsi basis	Nilai Fungsi basis	Parameter
Respon 1	B_{10}	1 (intersep)	17,08
	B_{11}	$(0,5307 - x_2)_+$	-10,24
	B_{12}	$(x_1 - 0,7230)_+$	56,14
	B_{13}	$(0,7230 - x_1)_+$	-28,46
	B_{14}	$(x_2 - 0,5307)_+(0,3155 - x_1)_+$	-85,83
	B_{15}	$(x_1 - 0,2671)_+$	-42,43
	B_{16}	$(0,1894 - x_1)_+(x_2 - 0,4450)_+$	84,66
	B_{17}	$(x_2 - 0,3654)_+$	-2,18
	B_{18}	$(0,5990 - x_1)_+$	10,30
Respon 2	B_{20}	1 (intersep)	-9,94
	B_{21}	$(0,8295 - x_3)_+$	24,97
	B_{22}	$(0,3850 - x_4)_+$	13,95
	B_{23}	$(0,8295 - x_3)_+(x_4 - 0,6599)_+$	15,75
	B_{24}	$(0,2509 - x_3)_+(x_4 - 0,5752)_+$	-51,06
	B_{25}	$(0,2509 - x_3)_+(0,5752 - x_4)_+$	-33,64
	B_{26}	$(0,3228 - x_3)_+$	-21,58
	B_{27}	$(x_3 - 0,6488)_+(0,6711 - x_4)_+$	41,79
	B_{28}	$(x_3 - 0,6488)_+(x_4 - 0,1277)_+$	23,69

Bentuk grafik estimasi fungsi regresi dari persamaan (5.8) disajikan dalam Gambar 5.24 masing-masing untuk respon 1 dan respon 2.

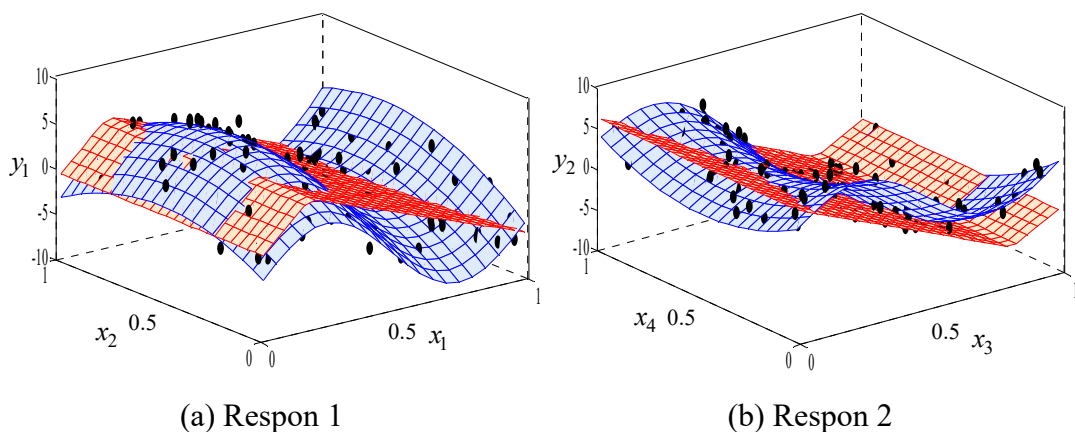


keterangan gambar: • observasi, fungsi asli dan estimasi fungsi regresi

Gambar 5.24. Grafik Estimasi Fungsi Regresi pada Simulasi
Fungsi Trigonometri-Polinomial dengan Fungsi Basis Optimal

Nilai $RMSE$ dan R^2 yang dihasilkan oleh estimasi model MARS birespon kontinu dalam persamaan (5.8) masing-masing sebesar 1,0223 dan 92,89%. Fungsi basis optimal MARS birespon kontinu dalam persamaan (5.8) dipilih dengan proses *stepwise* (*forward* dan *backward stepwise*) pada masing-masing respon secara parsial berdasarkan nilai GCV minimum. Dalam pemilihan fungsi basis optimal ini, diperoleh nilai GCV minimum untuk respon 1 sama dengan 1,5507 dan untuk respon 2 sama dengan 1,4386. Untuk mengakomodir korelasi antar respon, maka parameter regresi MARS birespon kontinu dalam persamaan (5.8) diperoleh dari hasil estimasi yang melibatkan matriks kovarians *error* (matriks bobot) dengan fungsi basis optimal pada masing-masing respon, dengan menggunakan persamaan (4.23). Sebagai perbandingan antara estimasi fungsi regresi yang memuat fungsi basis optimal dengan fungsi basis tidak optimal, maka disajikan bentuk estimasi fungsi regresi dengan sedikit fungsi basis dan banyak fungsi basis masing-masing pada Gambar 5.25 dan 5.26.

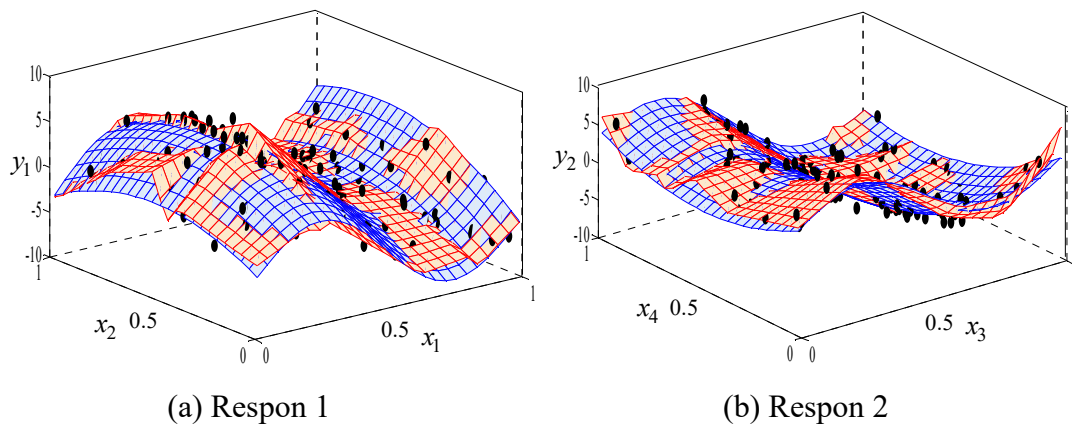
Gambar 5.25 menunjukkan bentuk estimasi fungsi regresi dengan sedikit fungsi basis (tidak optimal). Banyaknya fungsi basis untuk respon 1 adalah 3 fungsi basis dengan nilai GCV sebesar 5,3515 dan untuk respon 2 sebanyak 3 fungsi basis dengan nilai GCV sebesar 5,6793. Nilai $RMSE$ dan R^2 dari estimasi model MARS birespon ini masing-masing sebesar 2,2358 dan 65,97%. Dibandingkan dengan nilai $RMSE$ dan R^2 dari model optimal, tampak bahwa model ini memiliki keakuratan yang lebih rendah.



keterangan gambar: • observasi, ■ fungsi asli dan ■ estimasi fungsi regresi

Gambar 5.25. Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Trigonometri-Polinomial dengan Sedikit Fungsi Basis

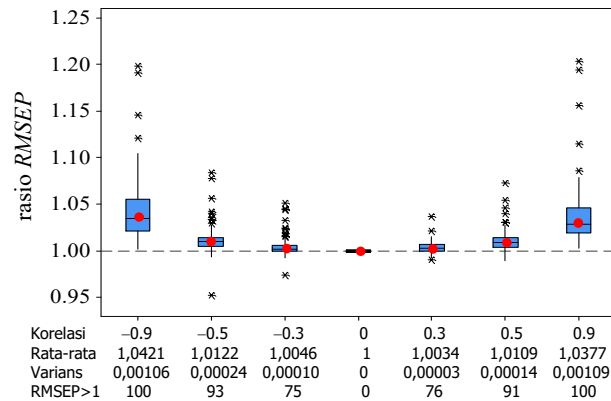
Gambar 5.26 menunjukkan bentuk estimasi fungsi regresi dengan banyak fungsi basis (tidak optimal). Banyaknya fungsi basis untuk respon 1 adalah 35 fungsi basis dengan nilai GCV sebesar 6,7795 dan untuk respon 2 sebanyak 35 fungsi basis dengan nilai GCV sebesar 6,6556. Nilai $RMSE$ dan R^2 dari estimasi model MARS birespon ini masing-masing sebesar 0,8236 dan 95,38%. Dibandingkan dengan nilai $RMSE$ dan R^2 dari model optimal, tampak bahwa model ini memiliki keakuratan yang lebih tinggi, penyebabnya karena fungsi basis yang termuat dalam model sangat banyak. Hal ini menyebabkan model yang dihasilkan sangat kompleks (tidak parsimoni) dan memungkinkan kasus *overfitting* dalam regresi.



keterangan gambar: • observasi, ■ fungsi asli dan ■ estimasi fungsi regresi

Gambar 5.26. Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Trigonometri-Polinomial dengan Banyak Fungsi Basis

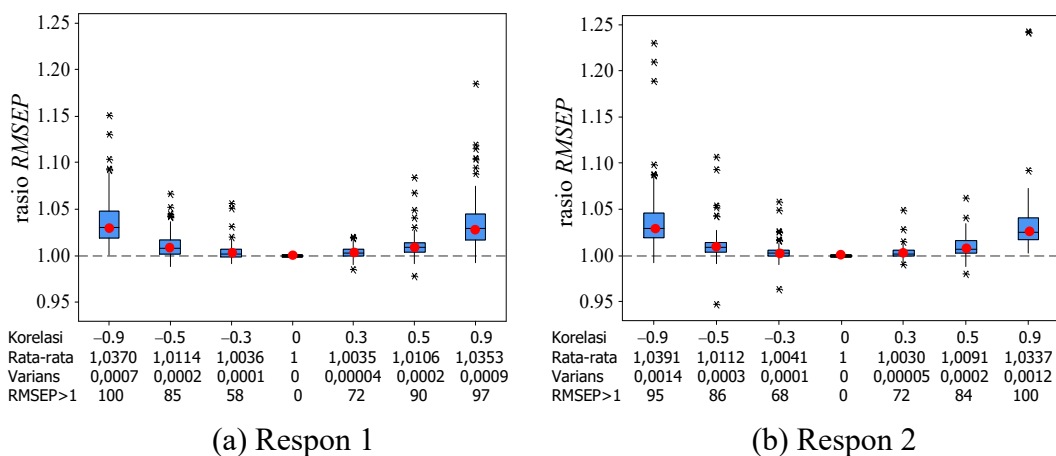
Untuk menunjukkan pengaruh korelasi terhadap model MARS birespon kontinu, maka dihitung rasio $RMSEP$ yang membandingkan antara $RMSEP$ model yang tidak melibatkan korelasi dengan $RMSEP$ model yang melibatkan korelasi. Gambar 5.27 menyajikan contoh perbandingan nilai rasio $RMSEP$ birespon untuk berbagai ukuran korelasi ($\rho = -0,9, -0,5, -0,3, 0, 0,3, 0,5, 0,9$) pada model persamaan (5.7) dengan replikasi sebanyak 100 kali percobaan. Agar lebih mudah dalam membandingkan pengaruh korelasi terhadap model MARS birespon kontinu, maka dalam simulasi ini ditetapkan $n_1 = n_2$ dan $\sigma_{11} = \sigma_{22}$, dimana pada simulasi ini diuraikan salah satu contoh untuk $n_1 = n_2 = 125$ dan $\sigma_{11} = \sigma_{22} = 1$.



Gambar 5.27. *Boxplot* Rasio *RMSEP* Birespon berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Trigonometri-Polinomial

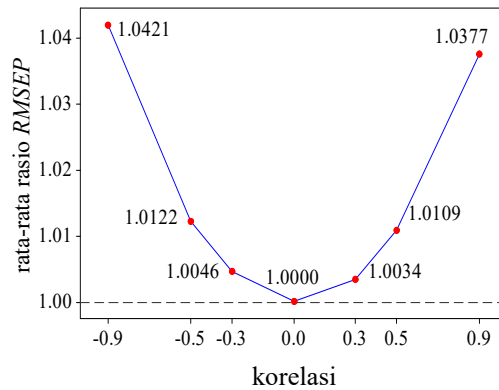
Gambar 5.27 menunjukkan bahwa semakin meningkat nilai mutlak korelasi ($|\rho|$), maka semakin besar penyebaran (varians) nilai rasio *RMSEP* birespon yang dihasilkan. Nilai rasio *RMSEP* birespon menurut gambar tersebut umumnya berada diatas nilai 1 seiring dengan meningkatnya nilai mutlak korelasi. Hal ini berarti bahwa nilai *RMSEP* model yang melibatkan korelasi umumnya menghasilkan nilai *RMSEP* yang lebih kecil daripada tidak melibatkan korelasi, seiring dengan meningkatnya nilai mutlak korelasi.

Apabila pemodelan birespon diuraikan untuk setiap respon, ternyata nilai rasio *RMSEP* untuk respon 1 dan respon 2 pada Gambar 5.28, menunjukkan pola yang sama dengan nilai rasio *RMSEP* birespon pada Gambar 5.27. Hal ini memperjelas bukti bahwa korelasi memberikan pengaruh pada penurunan *RMSEP*.



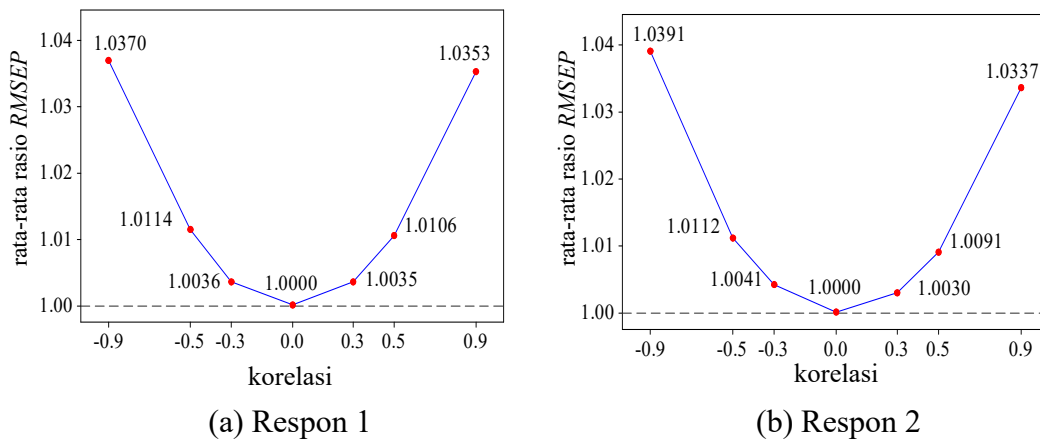
Gambar 5.28. *Boxplot* Rasio *RMSEP* berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Trigonometri-Polinomial pada Setiap Respon

Selanjutnya berdasarkan data *boxplot* pada Gambar 5.27, maka dapat disajikan perbandingan rata-rata nilai rasio *RMSEP* birespon pada berbagai ukuran korelasi dalam Gambar 5.29.



Gambar 5.29. Rata-rata Rasio *RMSEP* Birespon berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Trigonometri-Polinomial

Perbandingan rata-rata nilai rasio *RMSEP* pada berbagai ukuran korelasi juga dapat diuraikan untuk setiap respon berdasarkan data *boxplot* Gambar 5.28, yang hasilnya disajikan pada Gambar 5.30.



Gambar 5.30. Rata-rata Rasio *RMSEP* berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Trigonometri-Polinomial pada Setiap Respon

Gambar 5.29 dan 5.30 secara jelas menunjukkan pengaruh variasi besaran ukuran korelasi pada rata-rata nilai *RMSEP* model MARS birespon kontinu dalam simulasi fungsi trigonometri-polinomial. Hasil simulasi pada gambar tersebut menunjukkan bahwa semakin meningkat nilai mutlak korelasi pada model birespon, maka semakin besar rata-rata nilai rasio *RMSEP* yang

dihasilkan. Hal ini menjelaskan bahwa semakin besar nilai mutlak korelasi, maka semakin kecil nilai *RMSEP* dari model yang melibatkan korelasi dibandingkan dengan tidak melibatkan korelasi, berarti hasil prediksinya semakin baik. Rata-rata nilai rasio *RMSEP* juga menunjukkan bahwa nilai mutlak korelasi diatas 0,3 cenderung memberikan hasil prediksi MARS birespon kontinu yang lebih baik.

5.2.2. Simulasi MARS Birespon pada Fungsi Eksponensial-Trigonometri

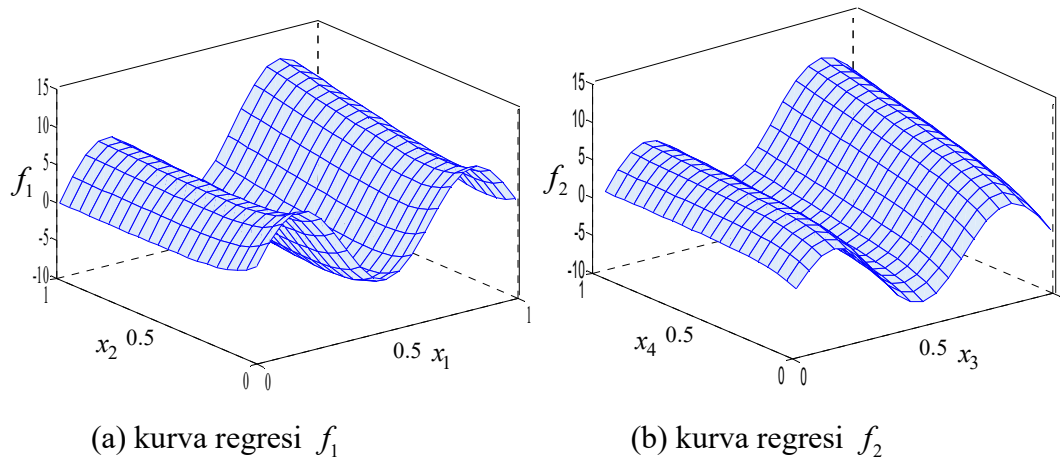
Sub bab berikut ini menguraikan hasil simulasi MARS birespon kontinu pada kombinasi fungsi eksponensial dan trigonometri. Simulasi dilakukan untuk mengetahui kemampuan MARS birespon kontinu dalam mendapatkan hasil estimasi yang mendekati bentuk fungsi eksponensial-trigonometri melalui pendekatan grafik. Contoh persamaan regresi untuk simulasi ini sebagai berikut:

$$\begin{aligned} y_{1i} &= 6 \sin(3\pi x_{1i}) + 3 \exp(-9x_{2i}) + 7x_{1i}x_{2i} + \varepsilon_{1i}, \\ y_{2i} &= 5 \sin(3\pi x_{3i}) - 2 \exp(-8x_{4i}) + 8x_{3i}x_{4i} + \varepsilon_{2i}, \end{aligned} \quad (5.9)$$

dimana prediktor x_{ji} ($j=1,2,3,4$; $i=1,2,\dots,125$) dibangkitkan dari distribusi uniform: $x_{ji} \sim U(0,1)$, sedangkan *random error* $\varepsilon = (\varepsilon'_1, \varepsilon'_2)'$ dibangkitkan dari

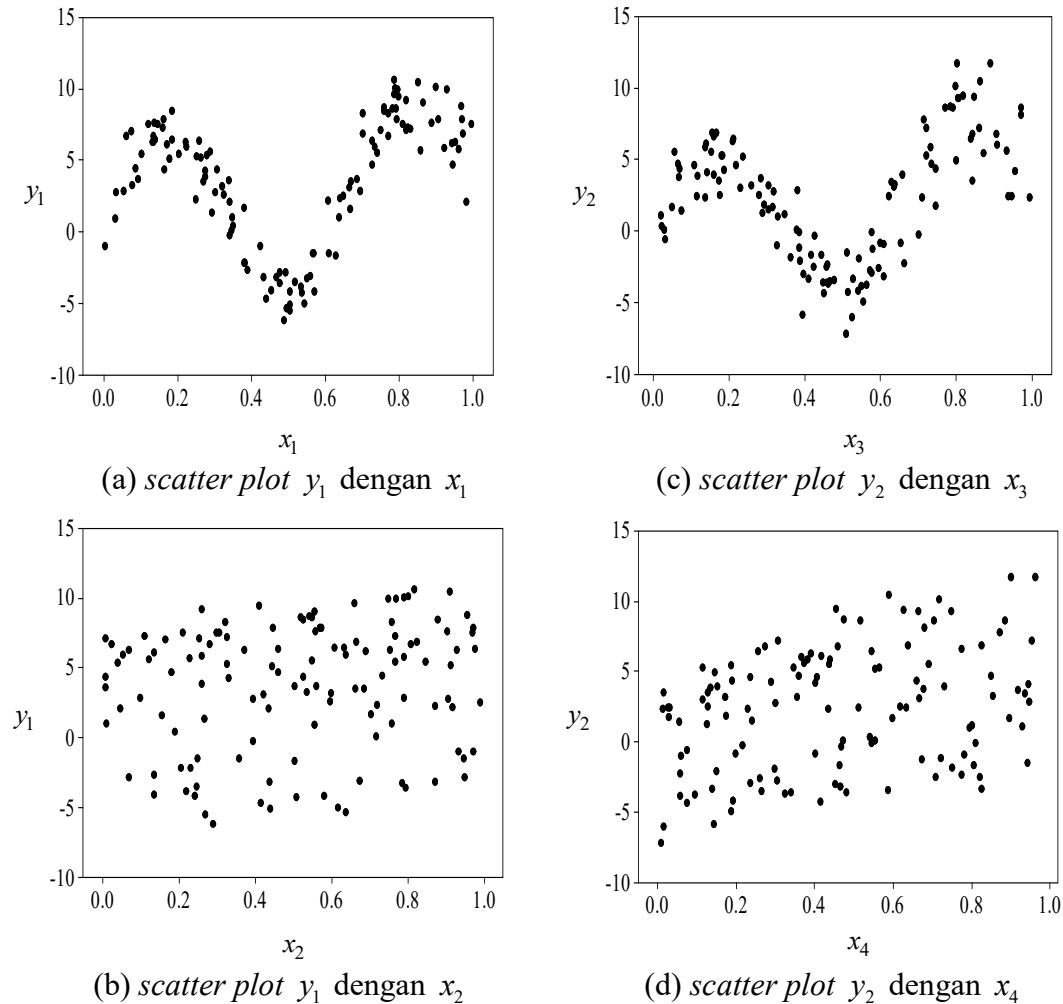
distribusi normal: $\varepsilon \sim N(0, \mathbf{W})$. Matriks $\mathbf{W} = \begin{pmatrix} \sigma_{11}\mathbf{I} & \sigma_{12}\mathbf{I} \\ \sigma_{12}\mathbf{I} & \sigma_{22}\mathbf{I} \end{pmatrix}$, dimana dalam

simulasi ini varians *error* $\sigma_{11} = \sigma_{22} = 1$, sedangkan kovarians *error* $\sigma_{12} = \rho\sqrt{\sigma_{11}\sigma_{22}}$ dengan korelasi *error* $\rho = 0,5$. Bentuk kurva regresi dalam persamaaan (5.9) disajikan dalam Gambar 5.31.



Gambar 5.31. Bentuk kurva regresi pada Simulasi Eksponensial-Trigonometri

Data hasil bangkitan prediktor dan respon berdasarkan simulasi fungsi percobaan ini disajikan dalam lampiran 12, dengan pola hubungan antara prediktor dengan respon disajikan dalam Gambar 5.32.



Gambar 5.32. *Scatter Plot* Respon dengan Prediktor pada Simulasi Fungsi Eksponensial-Trigonometri

Gambar 5.32 (a) dan (c) menunjukkan adanya perubahan perilaku pola hubungan antara respon dan prediktor dalam kurva regresi, selain itu Gambar 5.32 (b) dan (d) menunjukkan pola hubungan yang tidak diketahui bentuknya. Hal ini mengindikasikan penggunaan regresi nonparametrik salah satunya MARS birespon kontinu untuk mengestimasi model. Berdasarkan hasil pengolahan, diperoleh estimasi model (estimasi parameter dan fungsi basis optimal) pada simulasi Fungsi Percobaan Eksponensial-Trigonometri yang disajikan dalam Tabel 5.6.

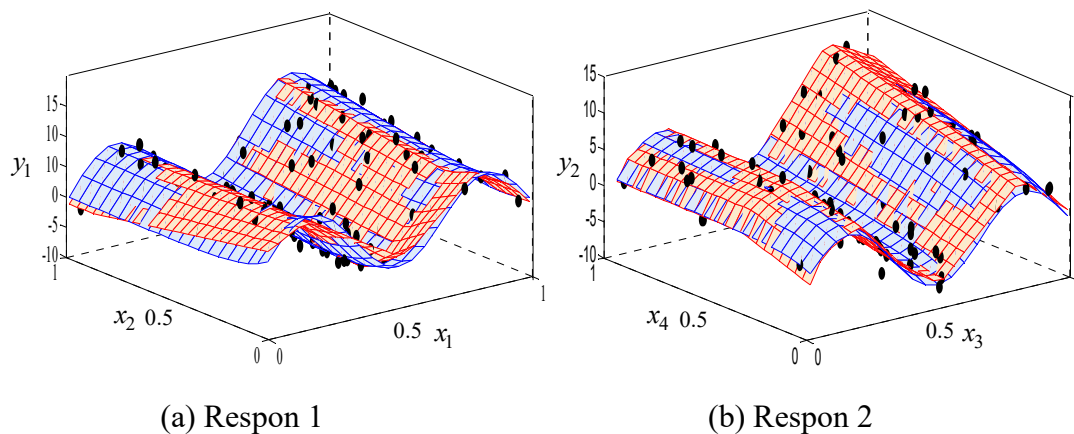
Tabel 5.6. Estimasi Parameter dan Fungsi Basis Optimal MARS
Birespon pada Simulasi Fungsi Eksponensial-Trigonometri

Respon	Fungsi basis	Nilai Fungsi basis	Parameter
Respon 1	B ₁₀	1 (intersep)	26,66
	B ₁₁	$(x_{1i} - 0,4906)_+$	56,36
	B ₁₂	$(x_{1i} - 0,7898)_+$	-68,20
	B ₁₃	$(x_{1i} - 0,1827)_+$	-68,41
	B ₁₄	$(x_{1i} - 0,1827)_+ (x_{2i} - 0,1555)_+$	7,66
	B ₁₅	$(x_{1i} - 0,1827)_+ (0,1555 - x_{2i})_+$	97,05
	B ₁₆	$(0,1827 - x_{1i})_+ (0,8757 - x_{2i})_+$	35,31
	B ₁₇	$(x_{1i} - 0,5426)_+$	36,94
	B ₁₈	$(0,1307 - x_{1i})_+$	-28,80
	B ₁₉	$(0,9288 - x_{1i})_+$	-25,86
	B ₁₀	$(x_{1i} - 0,4906)_+ (0,1555 - x_{2i})_+$	-155,1
Respon 2	B ₂₀	1	16,08
	B ₂₁	$(x_{3i} - 0,5140)_+$	-26,67
	B ₂₂	$(0,5140 - x_{3i})_+$	94,21
	B ₂₃	$(0,8167 - x_{3i})_+$	-69,57
	B ₂₄	$(x_{3i} - 0,1588)_+ (x_{4i} - 0,3461)_+$	12,21
	B ₂₅	$(x_{3i} - 0,1588)_+ (0,3461 - x_{4i})_+$	-13,85
	B ₂₆	$(0,5140 - x_{3i})_+ (0,3461 - x_{4i})_+$	-20,53
	B ₂₇	$(0,0673 - x_{3i})_+$	-53,24
	B ₂₈	$(0,2078 - x_{3i})_+$	-59,89
	B ₂₉	$(0,3966 - x_{3i})_+$	19,89

Persamaan regresi MARS birespon kontinu yang diperoleh dari estimasi parameter dan fungsi basis optimal dalam Tabel 5.6 dinyatakan sebagai berikut:

$$\begin{aligned}
 \hat{f}_1(x_{1i}, x_{2i}) = & 26,66 + 56,36(x_{1i} - 0,49)_+ - 68,20(x_{1i} - 0,79)_+ - 68,41(x_{1i} - 0,18)_+ + \\
 & 7,66(x_{1i} - 0,18)_+ (x_{2i} - 0,16)_+ + 97,05(x_{1i} - 0,18)_+ (0,16 - x_{2i})_+ + \\
 & 35,31(0,18 - x_{1i})_+ (0,88 - x_{2i})_+ + 36,94(x_{1i} - 0,54)_+ - 28,8(0,13 - x_{1i})_+ - \\
 & 25,86(0,93 - x_{1i})_+ - 155,1(x_{1i} - 0,49)_+ (0,16 - x_{2i})_+, \\
 \hat{f}_2(x_{3i}, x_{4i}) = & 16,08 - 26,67(x_{3i} - 0,51)_+ + 94,21(0,51 - x_{3i})_+ - 69,57(0,82 - x_{3i})_+ + \\
 & 12,21(x_{3i} - 0,16)_+ (x_{4i} - 0,35)_+ - 13,85(x_{3i} - 0,16)_+ (0,35 - x_{4i})_+ - \\
 & 20,53(0,51 - x_{3i})_+ (0,35 - x_{4i})_+ - 53,24(0,07 - x_{3i})_+ - 59,89(0,21 - x_{3i})_+ \\
 & + 19,89(0,40 - x_{3i})_+.
 \end{aligned}
 \tag{5.10}$$

Nilai $RMSE$ dan R^2 yang dihasilkan oleh model MARS birespon kontinu pada simulasi fungsi percobaan eksponensial-trigonometri dalam persamaan (5.10) masing-masing sebesar 1,0184 dan 94,57%. Bentuk grafik estimasi fungsi regresi dalam persamaan (5.10) disajikan dalam Gambar 5.33, masing-masing untuk respon 1 dan respon 2.

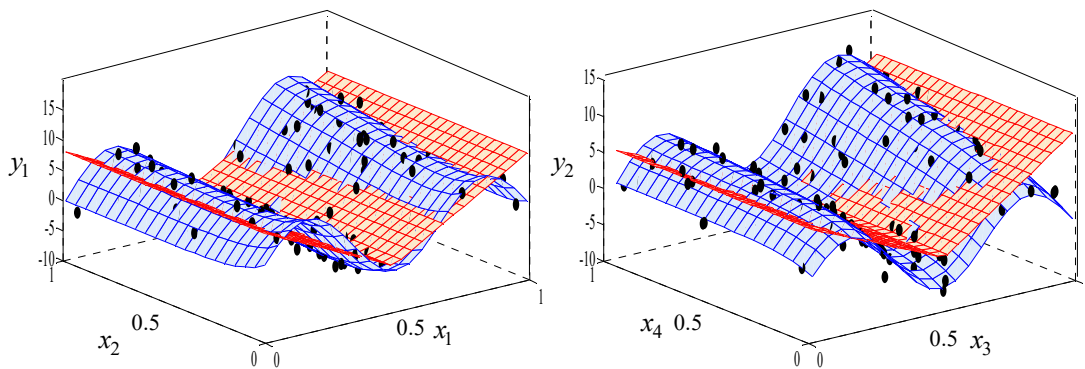


keterangan gambar: • observasi, ■ fungsi asli dan ■ estimasi fungsi regresi

Gambar 5.33. Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Eksponensial-Trigonometri dengan Fungsi Basis Optimal

Fungsi basis optimal MARS birespon kontinu dalam persamaan (5.10) dipilih dengan proses *stepwise* (*forward* dan *backward stepwise*) pada masing-masing respon secara parsial berdasarkan nilai GCV minimum. Dalam pemilihan fungsi basis optimal ini, diperoleh nilai GCV minimum untuk respon 1 sama dengan 1,6779 dan untuk respon 2 sama dengan 1,5074. Untuk mengakomodir korelasi antar respon, maka parameter regresi MARS birespon kontinu dalam persamaan (5.10) diperoleh dari hasil estimasi yang melibatkan matriks kovarians *error* (matriks bobot) dengan fungsi basis optimal pada masing-masing respon, dengan menggunakan persamaan (4.23). Sebagai perbandingan antara estimasi fungsi regresi yang memuat fungsi basis optimal dengan fungsi basis tidak optimal, maka disajikan bentuk estimasi fungsi regresi MARS birespon kontinu dengan sedikit fungsi basis dan banyak fungsi basis masing-masing pada Gambar 5.34 dan Gambar 5.35.

Gambar 5.34 menunjukkan bentuk estimasi fungsi regresi dengan sedikit fungsi basis (tidak optimal). Banyaknya fungsi basis untuk respon 1 adalah 3 fungsi basis dengan nilai GCV sebesar 8,5763 dan untuk respon 2 sebanyak 3 fungsi basis dengan nilai GCV sebesar 8,6479. Nilai $RMSE$ dan R^2 dari estimasi model MARS birespon ini masing-masing sebesar 2,7944 dan 59,14%. Dibandingkan dengan nilai $RMSE$ dan R^2 dari model optimal, tampak bahwa model ini memiliki keakuratan yang rendah.



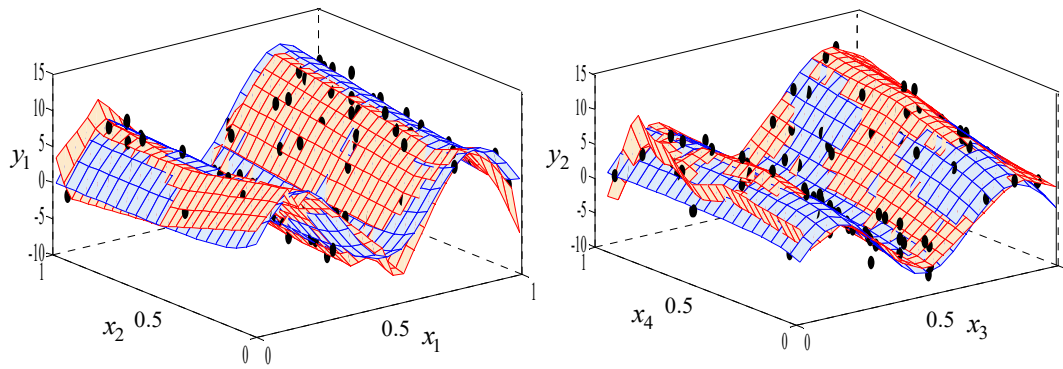
(a) Respon 1

(b) Respon 2

keterangan gambar: • observasi, ■ fungsi asli dan ■ estimasi fungsi regresi

Gambar 5.34. Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Eksponensial-Trigonometri dengan sedikit Fungsi Basis

Gambar 5.35 menunjukkan bentuk estimasi fungsi regresi dengan banyak fungsi basis (tidak optimal). Banyaknya fungsi basis untuk respon 1 adalah 35 fungsi basis dengan nilai GCV sebesar 7,9767 dan untuk respon 2 sebanyak 31 fungsi basis dengan nilai GCV sebesar 4,5439. Nilai $RMSE$ dan R^2 dari estimasi model MARS birespon ini masing-masing sebesar 0,8759 dan 95,99%. Dibandingkan dengan nilai $RMSE$ dan R^2 dari model optimal, tampak bahwa model ini memiliki keakuratan yang lebih tinggi, penyebabnya karena fungsi basis yang termuat dalam model sangat banyak. Hal ini menyebabkan model yang dihasilkan sangat kompleks (tidak parsimoni), yang memungkinkan munculnya kasus *overfitting*.



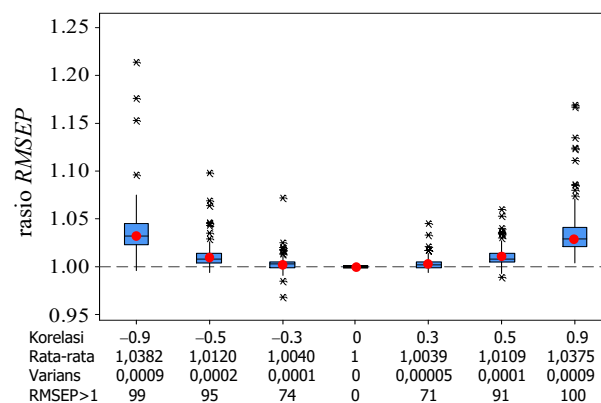
(a) Respon 1

(b) Respon 2

keterangan gambar: • observasi, ■ fungsi asli dan ■ estimasi fungsi regresi

Gambar 5.35. Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Eksponensial-Trigonometri dengan Banyak Fungsi Basis

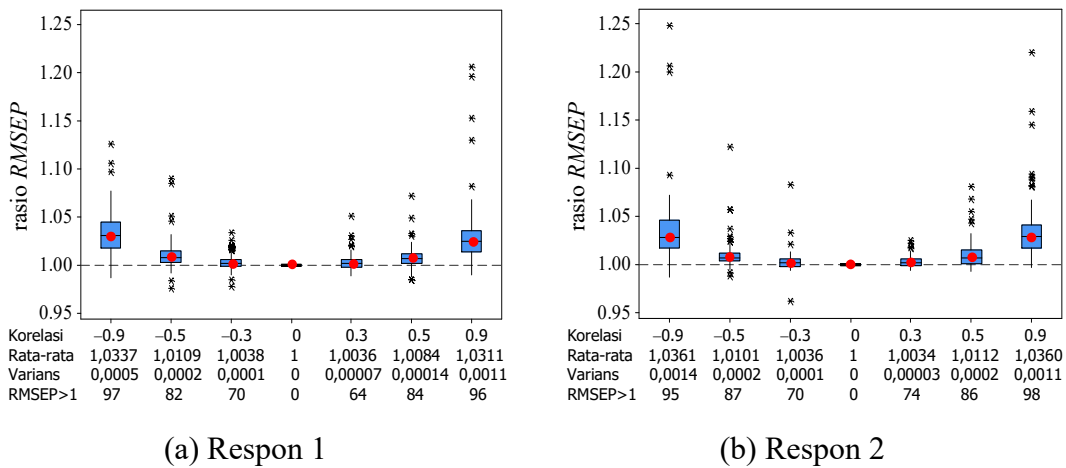
Untuk menunjukkan pengaruh korelasi terhadap model MARS birespon kontinu, maka dalam simulasi fungsi eksponensial-trigonometri ini dihitung rasio *RMSEP* yang membandingkan antara *RMSEP* model yang tidak melibatkan korelasi dengan *RMSEP* model yang melibatkan korelasi. Gambar 5.36 menyajikan perbandingan nilai rasio *RMSEP* birespon untuk berbagai korelasi ($\rho = -0,9, -0,5, -0,3, 0, 0,3, 0,5, 0,9$) pada model persamaan (5.9) dengan replikasi sebanyak 100 kali. Untuk melihat pengaruh korelasi, maka dalam simulasi ini ditetapkan $n_1 = n_2$ dan $\sigma_{11} = \sigma_{22}$, dimana pada simulasi ini diuraikan salah satu contoh untuk $n_1 = n_2 = 125$ dan $\sigma_{11} = \sigma_{22} = 1$.



Gambar 5.36. Boxplot Rasio *RMSEP* Birespon berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Eksponensial-Trigonometri

Gambar 5.36 menunjukkan bahwa semakin meningkat nilai mutlak korelasi ($|\rho|$), maka semakin besar penyebaran (varians) nilai rasio *RMSEP* birespon yang dihasilkan. Nilai rasio *RMSEP* birespon menurut gambar tersebut umumnya berada diatas nilai 1 seiring dengan meningkatnya nilai mutlak korelasi. Hal ini berarti bahwa nilai *RMSEP* dari model yang melibatkan korelasi umumnya menghasilkan nilai *RMSEP* model yang lebih kecil daripada tidak melibatkan korelasi, seiring dengan meningkatnya nilai mutlak korelasi.

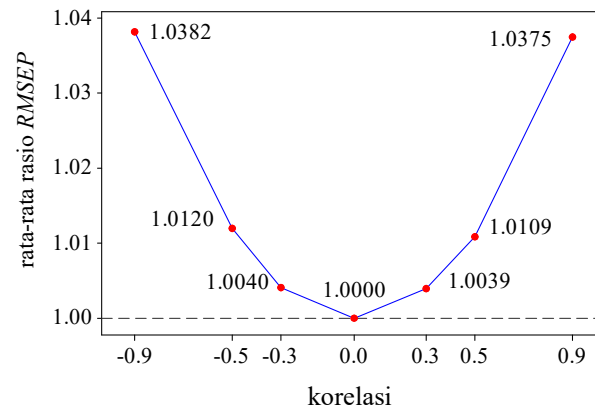
Apabila hasil pemodelan MARS birespon kontinu diuraikan untuk setiap respon, ternyata nilai rasio *RMSEP* untuk respon 1 dan respon 2 dalam Gambar 5.37 menunjukkan pola yang sama dengan nilai rasio *RMSEP* birespon dalam Gambar 5.36. Hal ini memperjelas bukti bahwa besarnya nilai mutlak korelasi memberikan pengaruh pada penurunan nilai *RMSEP* birespon yang dihasilkan.



Gambar 5.37 *Boxplot* Rasio *RMSEP* berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Eksponensial-Trigonometri pada Setiap Respon

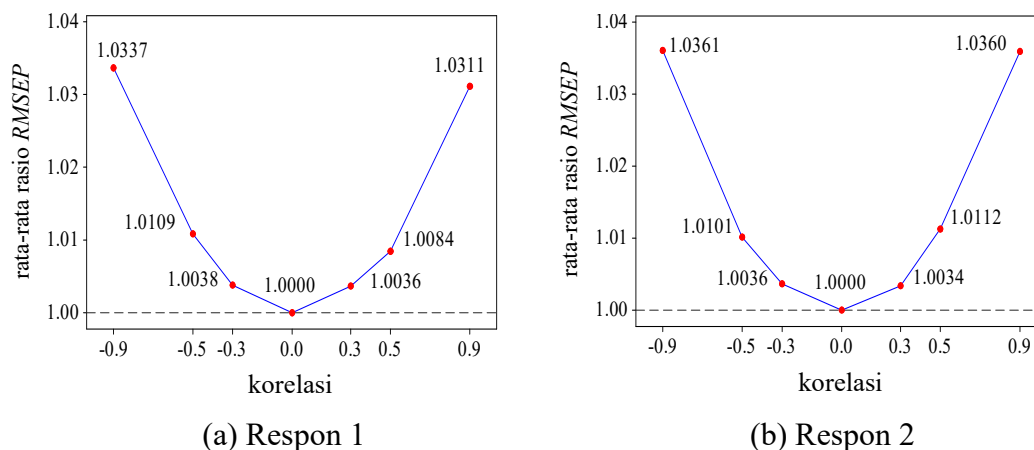
Selanjutnya berdasarkan data *boxplot* pada Gambar 5.36, maka dapat disajikan perbandingan rata-rata rasio *RMSEP* birespon pada berbagai ukuran korelasi dalam Gambar 5.38. Hasil perbandingan secara jelas menunjukkan pengaruh variasi besaran ukuran korelasi pada rata-rata nilai *RMSEP* model MARS birespon kontinu dalam simulasi fungsi trigonometri-polinomial. Hasil simulasi pada gambar tersebut menunjukkan bahwa semakin meningkat nilai mutlak korelasi pada model birespon, maka semakin besar rata-rata nilai rasio

$RMSEP$ birespon yang dihasilkan. Hal ini menjelaskan bahwa semakin besar nilai mutlak korelasi dalam model MARS birespon kontinu, maka semakin kecil nilai $RMSEP$ dari model yang melibatkan korelasi dibandingkan dengan model yang tidak melibatkan korelasi.



Gambar 5.38. Rata-Rata Rasio $RMSEP$ Birespon berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Eksponensial-Trigonometri

Perbandingan rata-rata nilai rasio $RMSEP$ pada berbagai ukuran korelasi juga dapat diuraikan untuk setiap respon berdasarkan data *boxplot* Gambar 5.37, yang hasilnya disajikan pada Gambar 5.39.



Gambar 5.39. Rata-Rata Rasio $RMSEP$ berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Eksponensial-Trigonometri pada Setiap Respon

Pola rata-rata rasio $RMSEP$ pada Gambar 5.39 juga menunjukkan pola yang sama dengan Gambar 5.38, yaitu bahwa semakin besar nilai mutlak korelasi antar

respon maka semakin baik penggunaan estimator MARS birespon kontinu, yang ditandai oleh semakin rendahnya nilai *RMSEP* model yang diperoleh. Rata-rata nilai rasio *RMSEP* juga menunjukkan bahwa nilai mutlak korelasi diatas 0,3 cenderung memberikan hasil prediksi MARS birespon kontinu yang lebih baik.

5.2.3. Simulasi MARS Birespon pada Fungsi Polinomial-Eksponensial

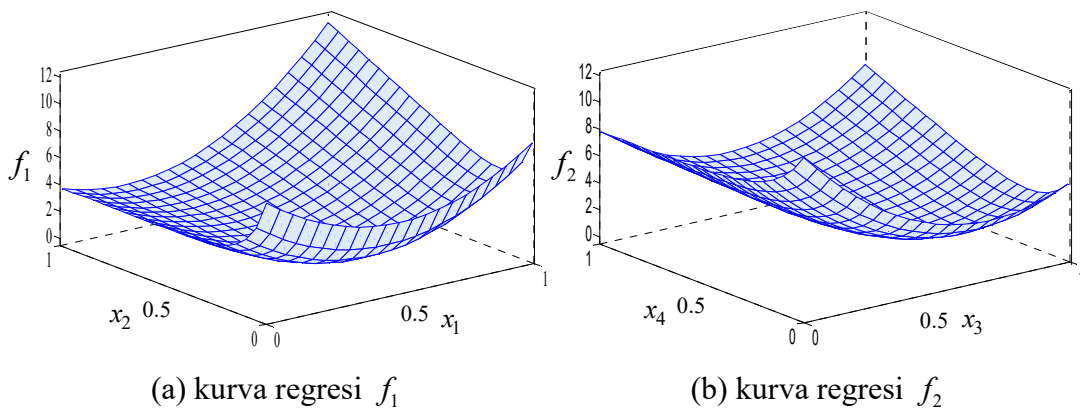
Subbab berikut ini menguraikan hasil simulasi MARS birespon kontinu pada kombinasi fungsi polinomial dan eksponensial. Simulasi dilakukan untuk mengetahui kemampuan MARS birespon kontinu dalam mendapatkan hasil estimasi yang mendekati bentuk fungsi polinomial-eksponensial melalui pendekatan grafik. Contoh persamaan regresi yang dirancang untuk studi simulasi ini adalah sebagai berikut:

$$\begin{aligned} y_{1i} &= 15(0,5 - x_{1i})^2 + 5 \exp(-12x_{2i}) + 8x_{1i}x_{2i} + \varepsilon_{1i}, \\ y_{2i} &= 16(x_{3i} - 0,7)^2 + 4 \exp(-7x_{4i}) + 7x_{3i}x_{4i} + \varepsilon_{2i}, \end{aligned} \quad (5.11)$$

dimana prediktor x_{ji} ($j = 1, 2, 3, 4; i = 1, 2, \dots, 125$) dibangkitkan dari distribusi uniform: $x_{ji} \sim U(0,1)$, sedangkan *random error* $\varepsilon = (\varepsilon'_1, \varepsilon'_2)'$ dibangkitkan dari

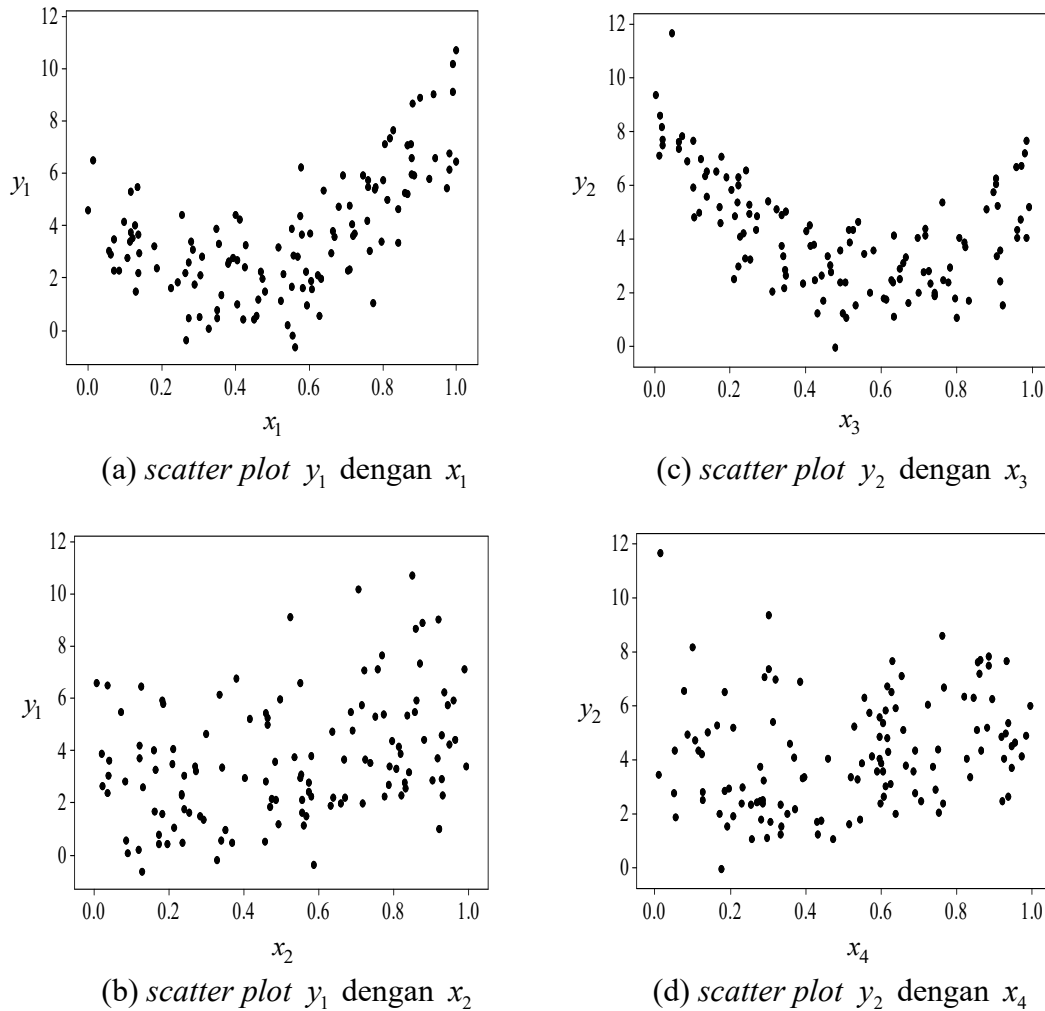
distribusi normal: $\varepsilon \sim N(0, \mathbf{W})$. Matriks $\mathbf{W} = \begin{pmatrix} \sigma_{11}\mathbf{I} & \sigma_{12}\mathbf{I} \\ \sigma_{12}\mathbf{I} & \sigma_{22}\mathbf{I} \end{pmatrix}$, dimana varians *error*

$\sigma_{11} = \sigma_{22} = 1$, sedangkan kovarians *error* $\sigma_{12} = \rho\sqrt{\sigma_{11}\sigma_{22}}$ dengan korelasi antar *error* $\rho = 0,5$. Bentuk kurva regresi dari persamaan (5.11) disajikan dalam Gambar 5.40, masing-masing untuk respon 1 dan respon 2.



Gambar 5.40. Bentuk kurva regresi pada Simulasi Fungsi Polinomial-Eksponensial

Hasil bangkitan prediktor dan respon berdasarkan fungsi percobaan ini disajikan pada data dalam lampiran 13, dengan pola hubungan antara prediktor dengan respon disajikan dalam Gambar 5.41.



Gambar 5.41. *Scatter Plot* Respon dengan Prediktor pada Simulasi Fungsi Polinomial-Eksponensial

Gambar 5.41 (a) dan (c) menunjukkan adanya perubahan perilaku pola hubungan antara respon dan prediktor dalam kurva regresi, selain itu Gambar 5.41 (b) dan (d) menunjukkan pola hubungan yang tidak diketahui bentuknya. Hal ini mengindikasikan penggunaan regresi nonparametrik salah satunya MARS birespon kontinu untuk estimasi model. Berdasarkan hasil pengolahan, diperoleh estimasi model (estimasi parameter dan fungsi basis optimal) dari simulasi fungsi percobaan Polinomial-Eksponensial yang disajikan dalam Tabel 5.7.

Tabel 5.7. Estimasi Parameter dan Fungsi Basis Optimal MARS
Birespon pada Simulasi Fungsi Polinomial-Ekspensial

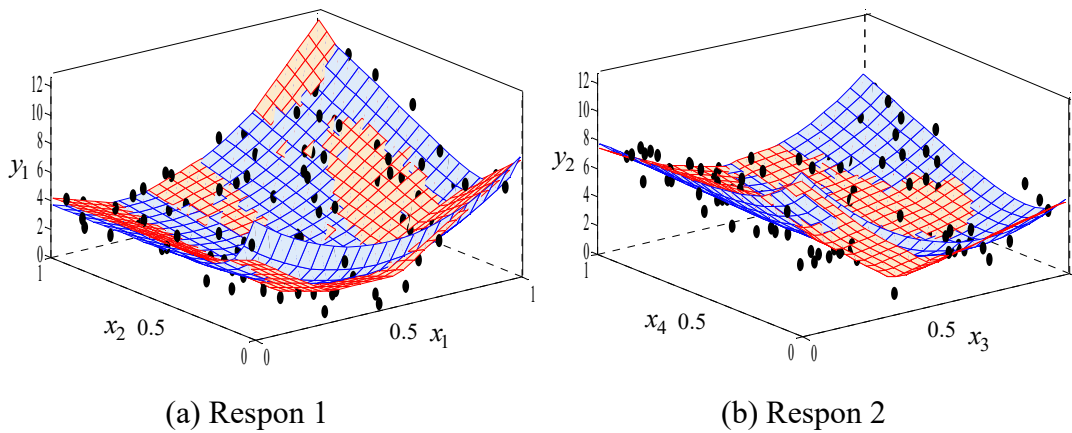
Respon	Fungsi basis	Nilai Fungsi basis	Parameter
Respon1	B ₁₀	1 (intersep)	0,83
	B ₁₁	$(x_{2i} - 0,2365)_+$	5,74
	B ₁₂	$(0,2365 - x_{2i})_+$	5,83
	B ₁₃	$(x_{2i} - 0,2635)_+ (x_{1i} - 0,5801)_+$	-159,2
	B ₁₄	$(0,2736 - x_{1i})_+$	12,65
	B ₁₅	$(x_{2i} - 0,2365)_+ (x_{1i} - 0,5942)_+$	188,1
	B ₁₆	$(x_{2i} - 0,2365)_+ (0,6608 - x_{1i})_+$	-8,75
	B ₁₇	$(x_{1i} - 0,5408)_+ (0,8608 - x_{2i})_+$	15,36
Respon 2	B ₂₀	1 (intersep)	-0,36
	B ₂₁	$(0,4910 - x_{3i})_+$	18,63
	B ₂₂	$(0,1923 - x_{4i})_+$	11,91
	B ₂₃	$(x_{3i} - 0,2505)_+$	4,33
	B ₂₄	$(0,8069 - x_{3i})_+ (x_{4i} - 0,1974)_+$	15,13
	B ₂₅	$(x_{4i} - 0,1923)_+ (x_{3i} - 0,5157)_+$	12,66
	B ₂₆	$(x_{4i} - 0,1923)_+ (0,5157 - x_{3i})_+$	-26,72

Persamaan regresi MARS birespon kontinu yang diperoleh dari estimasi parameter dan fungsi basis optimal dalam Tabel 5.7 dapat dinyatakan dalam persamaan sebagai berikut:

$$\begin{aligned}
 \hat{f}_1(x_{1i}, x_{2i}) = & 0,83 - 5,74(x_{2i} - 0,24)_+ + 5,83(0,24 - x_{2i})_+ - (159,2(x_{2i} - 0,26)_+ \times \\
 & (x_{1i} - 0,58)_+) + 12,65(0,27 - x_{1i})_+ + 188,1(x_{2i} - 0,24)_+ (x_{1i} - 0,59)_+ - \\
 & 8,75(x_{2i} - 0,24)_+ (0,66 - x_{1i})_+ + 15,36(x_{1i} - 0,54)_+ (0,86 - x_{2i})_+, \\
 \hat{f}_2(x_{3i}, x_{4i}) = & -0,36 + 18,63(0,49 - x_{3i})_+ + 11,91(0,19 - x_{4i})_+ + 4,33(x_{3i} - 0,25)_+ + \\
 & 15,13(0,81 - x_{3i})_+ (x_{4i} - 0,20)_+ + 12,66(x_{4i} - 0,19)_+ (x_{3i} - 0,52)_+ - \\
 & 26,72(x_{4i} - 0,19)_+ (0,52 - x_{3i})_+.
 \end{aligned}
 \tag{5.12}$$

Nilai $RMSE$ dan R^2 yang dihasilkan oleh model MARS birespon kontinu dalam persamaan (5.12) masing-masing sebesar 0,9734 dan 80,25%. Bentuk grafik

estimasi fungsi regresi dalam persamaan (5.12) disajikan dalam Gambar 5.42, masing-masing untuk respon 1 dan respon 2.



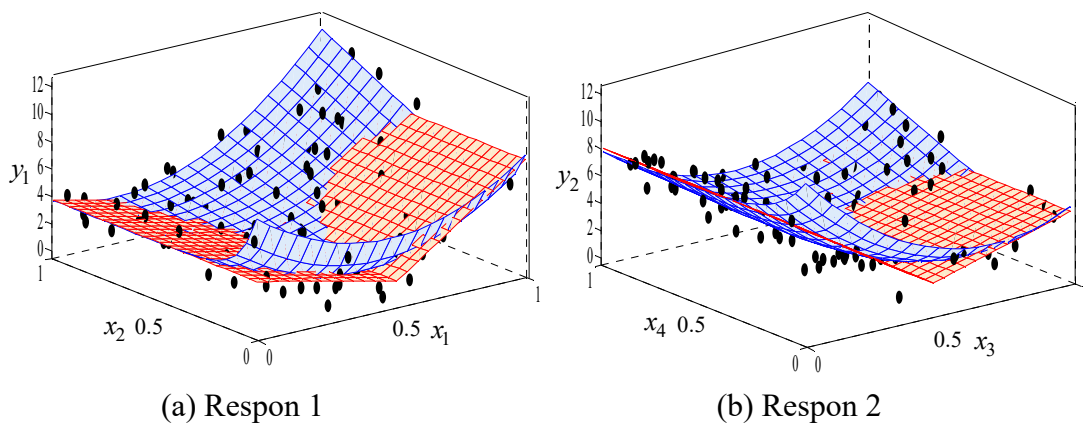
keterangan gambar: • observasi, ■ fungsi asli dan ■ estimasi fungsi regresi

Gambar 5.42. Grafik Estimasi Fungsi Regresi pada Simulasi
Fungsi Polinomial-Ekspensial dengan Fungsi Basis Optimal

Fungsi basis optimal MARS birespon kontinu dalam persamaan (5.12) dipilih dengan proses *stepwise* (*forward* dan *backward stepwise*) pada masing-masing respon secara parsial berdasarkan nilai *GCV* minimum. Dalam pemilihan fungsi basis optimal ini, diperoleh nilai *GCV* minimum untuk respon 1 sama dengan 1,3096 dan untuk respon 2 sama dengan 1,2266. Untuk mengakomodir korelasi antar respon, maka parameter regresi MARS birespon kontinu dalam persamaan (5.12) diperoleh dari hasil estimasi yang melibatkan matriks kovarians *error* (matriks bobot) dengan fungsi basis optimal pada masing-masing respon, dengan menggunakan persamaan (4.23). Sebagai perbandingan antara estimasi fungsi regresi yang memuat fungsi basis optimal dengan fungsi basis tidak optimal, maka disajikan bentuk estimasi fungsi regresi dengan sedikit fungsi basis dan banyak fungsi basis masing-masing pada Gambar 5.43 dan 5.44.

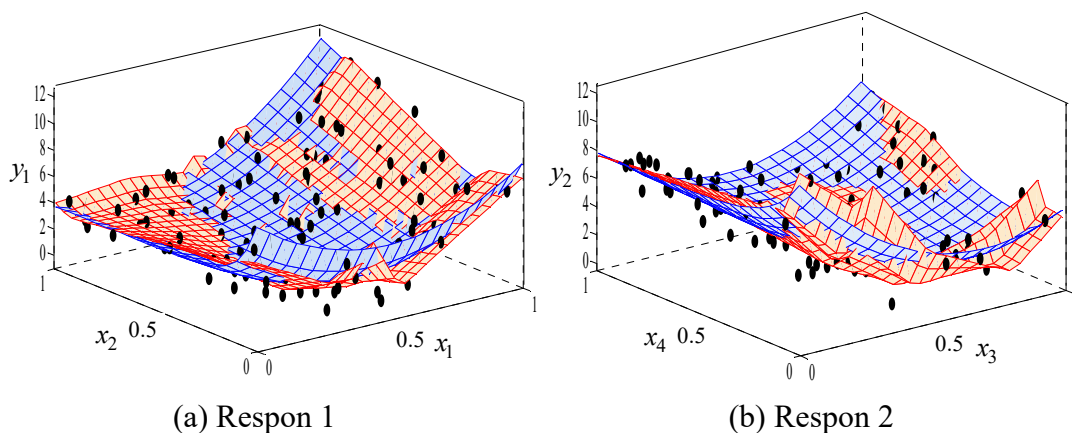
Gambar 5.43 menunjukkan bentuk estimasi fungsi regresi dengan sedikit fungsi basis (tidak optimal). Banyaknya fungsi basis untuk respon 1 adalah 3 fungsi basis dengan nilai *GCV* sebesar 2,2389 dan untuk respon 2 sebanyak 3 fungsi basis dengan nilai *GCV* sebesar 1,8189. Nilai *RMSE* dan R^2 dari estimasi model MARS birespon ini masing-masing sebesar 1,3561 dan 61,67%.

Dibandingkan dengan nilai $RMSE$ dan R^2 dari model optimal, tampak bahwa model ini memiliki keakuratan yang rendah.



keterangan gambar: • observasi, ■ fungsi asli dan ■ estimasi fungsi regresi

Gambar 5.43. Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Polinomial-Ekspensial dengan Sedikit Fungsi Basis



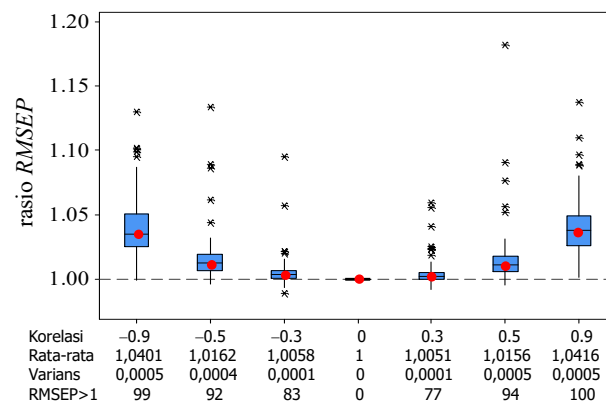
keterangan gambar: • observasi, ■ fungsi asli dan ■ estimasi fungsi regresi

Gambar 5.44. Grafik Estimasi Fungsi Regresi pada Simulasi Fungsi Polinomial-Ekspensial dengan Banyak Fungsi Basis

Gambar 5.44 menunjukkan bentuk estimasi fungsi regresi dengan banyak fungsi basis (tidak optimal). Banyaknya fungsi basis untuk respon 1 adalah 29 fungsi basis dengan nilai GCV sebesar 3,8598 dan untuk respon 2 sebanyak 27 fungsi basis dengan nilai GCV sebesar 3,2451. Nilai $RMSE$ dan R^2 dari estimasi model MARS birespon ini masing-masing sebesar 0,8589 dan 84,63%. Dibandingkan dengan nilai $RMSE$ dan R^2 dari model optimal, tampak bahwa

model ini memiliki keakuratan yang lebih tinggi, penyebabnya karena fungsi basis yang termuat dalam model sangat banyak. Hal ini menyebabkan model yang dihasilkan sangat kompleks (tidak parsimoni) dan memungkinkan kasus *overfitting* dalam regresi.

Untuk menunjukkan pengaruh korelasi terhadap model MARS birespon kontinu dalam simulasi fungsi polinomial-eksponensial ini, maka dihitung rasio *RMSEP* birespon yang membandingkan antara *RMSEP* model yang tidak melibatkan korelasi dengan *RMSEP* model yang melibatkan korelasi. Simulasi dilakukan dengan replikasi sebanyak 100 kali untuk melihat pola nilai rasio *RMSEP* yang dihasilkan oleh variasi besaran ukuran korelasi. Gambar 5.45 menyajikan perbandingan nilai rasio *RMSEP* birespon untuk berbagai korelasi ($\rho = -0,9, -0,5, -0,3, 0, 0,3, 0,5, 0,9$) pada model persamaan (5.11). Agar lebih mudah dalam membandingkan korelasi, maka dalam simulasi ini ditetapkan $n_1 = n_2$ dan $\sigma_{11} = \sigma_{22}$, dimana pada simulasi ini diuraikan salah satu contoh untuk $n_1 = n_2 = 125$ dan $\sigma_{11} = \sigma_{22} = 1$.

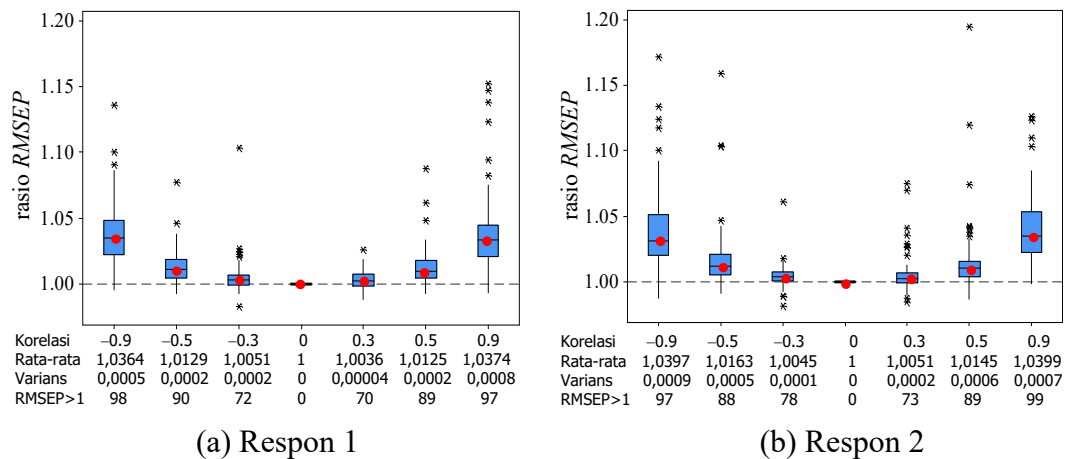


Gambar 5.45. *Boxplot* Rasio *RMSEP* Birespon berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Polinomial-Eksponensial

Gambar 5.45 menunjukkan bahwa semakin meningkat nilai mutlak korelasi ($|\rho|$), maka semakin besar penyebaran (varians) nilai rasio *RMSEP* birespon yang dihasilkan. Nilai rasio *RMSEP* birespon menurut gambar tersebut umumnya berada diatas nilai 1 seiring dengan meningkatnya nilai mutlak korelasi. Hal ini berarti bahwa nilai *RMSEP* birespon dari model yang melibatkan korelasi

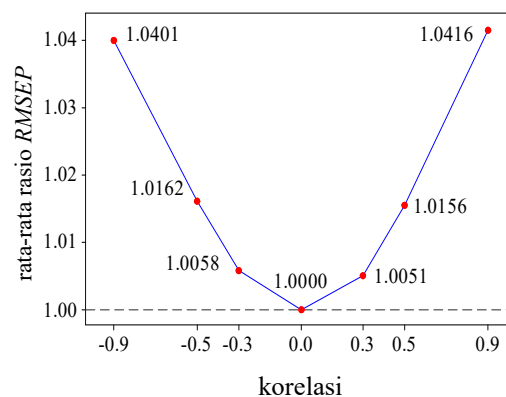
umumnya menghasilkan nilai $RMSEP$ model yang lebih kecil daripada tidak melibatkan korelasi, seiring dengan meningkatnya nilai mutlak korelasi.

Apabila hasil pemodelan MARS birespon kontinu diuraikan untuk setiap respon, ternyata nilai rasio $RMSEP$ untuk respon 1 dan respon 2 pada Gambar 5.46 menunjukkan pola yang sama dengan nilai rasio $RMSEP$ birespon yang disajikan pada Gambar 5.45. Hal ini memperjelas bukti bahwa besarnya nilai mutlak korelasi memberikan pengaruh pada penurunan nilai $RMSEP$ birespon.



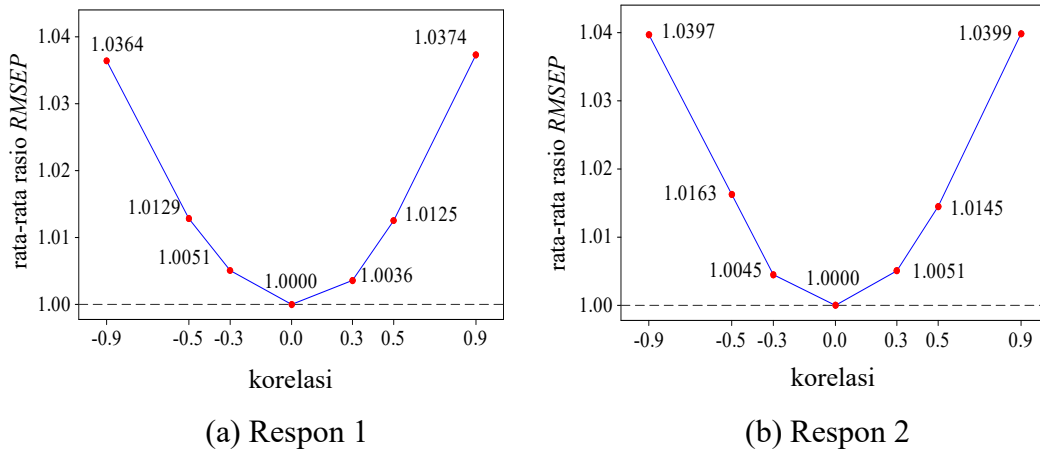
Gambar 5.46. *Boxplot* Rasio $RMSEP$ berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Polinomial-Ekspensial pada Setiap Respon

Selanjutnya berdasarkan data *boxplot* pada Gambar 5.45, maka dapat disajikan perbandingan rata-rata nilai rasio $RMSEP$ birespon pada berbagai ukuran korelasi dalam Gambar 5.47.



Gambar 5.47. Rata-Rata Rasio $RMSEP$ Birespon berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Polinomial-Ekspensial

Perbandingan rata-rata nilai rasio $RMSEP$ pada berbagai ukuran korelasi juga dapat diuraikan untuk setiap respon berdasarkan data *boxplot* Gambar 5.46, yang hasilnya disajikan pada Gambar 5.48.



Gambar 5.48. Rata-rata Rasio $RMSEP$ berdasarkan Variasi Ukuran Korelasi pada Simulasi Fungsi Polinomial-Eksponensial pada Setiap Respon

Gambar 5.47 dan 5.48 secara jelas menunjukkan pengaruh variasi besaran ukuran korelasi pada rata-rata nilai $RMSEP$ model MARS birespon kontinu dalam simulasi fungsi polinomial-eksponensial. Hasil simulasi pada gambar tersebut menunjukkan bahwa semakin meningkat nilai mutlak korelasi pada model birespon, maka semakin besar rata-rata nilai rasio $RMSEP$ birespon yang dihasilkan. Hal ini menjelaskan bahwa semakin besar nilai mutlak korelasi dalam model MARS birespon kontinu, maka semakin kecil nilai $RMSEP$ dari model yang melibatkan korelasi dibandingkan dengan tidak melibatkan korelasi, yang berarti hasil prediksinya semakin baik. Berdasarkan rata-rata nilai rasio $RMSEP$ juga menunjukkan bahwa nilai mutlak korelasi diatas 0,3 cenderung memberikan hasil prediksi MARS birespon kontinu yang lebih baik.

Berdasarkan uraian simulasi pada berbagai fungsi percobaan, maka dapat diambil beberapa kesimpulan sebagai berikut:

1. Estimator MARS birespon kontinu mampu digunakan untuk mengestimasi berbagai bentuk fungsi percobaan, seperti pada fungsi percobaan Trigonometri, Polinomial, dan Eksponensial, yang ditunjukkan melalui pendekatan grafik.

2. Jika fungsi basis yang digunakan sangat sedikit, maka tingkat keakuratan model akan sangat rendah, yang ditandai oleh tingginya nilai *RMSE* dan rendahnya nilai R^2 model. Sebaliknya jika fungsi basis yang digunakan sangat banyak, maka model yang dihasilkan akan sangat kompleks. Akibatnya perlu dipilih fungsi basis yang optimal yang paling sesuai untuk data.
3. Semakin besar nilai mutlak korelasi antar respon maka semakin baik penggunaan estimator MARS birespon, yang ditunjukkan oleh semakin rendahnya nilai *RMSEP* model yang diperoleh. Simulasi juga menunjukkan bahwa nilai mutlak korelasi diatas 0,3 ($|\rho| > 0,3$) cenderung memberikan hasil prediksi model MARS birespon kontinu yang lebih baik.

Halaman ini sengaja dikosongkan

BAB 6

PEMODELAN PERSENTASE PENDUDUK MISKIN DAN IPM DENGAN MARS BIRESPON KONTINU

Bab ini merupakan kajian empiris yang membahas hasil tujuan penelitian 4 yang disebutkan dalam Bab 1, yakni penerapan model MARS birespon kontinu pada data riil, yaitu pada pemodelan persentase penduduk miskin dan IPM di Pulau Jawa. Hasil kajian empiris pada bab ini diperoleh berdasarkan tahapan metodologi dalam Bab 3 (Sub Bab 3.4), yang hasilnya diuraikan dalam subbab-subbab di bawah ini. Hasil kajian diawali dengan eksplorasi data penelitian, dilanjutkan dengan pemodelan data dan memvalidasi model yang diperoleh.

6.1. Eksplorasi Data Persentase Penduduk Miskin, IPM dan Variabel yang Mempengaruhinya

Data riil yang digunakan dalam penelitian ini adalah data sekunder yang bersumber dari beberapa publikasi terbitan BPS tahun 2013, seperti publikasi INKESRA, IPM, PDRB, kemiskinan dan publikasi lainnya. Variabel yang digunakan dalam penelitian ini terdiri dari respon dan prediktor. Respon terdiri atas dua variabel dan saling berkorelasi, sedangkan prediktor terdiri atas lima variabel. Unit observasi adalah seluruh kabupaten/kota yang ada di Pulau Jawa (118 kabupaten/kota).

1. Respon (y) :

y_{1i} : persentase penduduk miskin pada wilayah ke- i ,

y_{2i} : IPM pada wilayah ke- i .

2. Prediktor (x) :

x_{1i} : pertumbuhan ekonomi pada wilayah ke- i ,

x_{2i} : angka beban ketergantungan pada wilayah ke- i ,

x_{3i} : persentase pengeluaran perkapita untuk makanan pada wilayah ke- i ,

x_{4i} : TPT pada wilayah ke- i ,

x_{5i} : persentase penduduk pekerjaan utama sektor pertanian pada wilayah ke- i .

Data masing-masing variabel penelitian disajikan pada lampiran 15. Untuk mendapatkan gambaran umum dari data yang digunakan, maka berikut ini disajikan tabel statistik deskriptif dari masing-masing variabel penelitian.

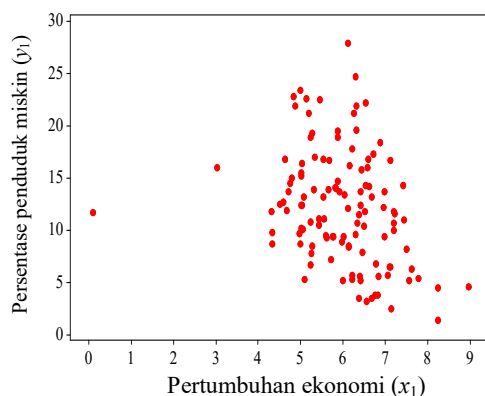
Tabel 6.1. Statistik Deskriptif Variabel Penelitian pada Data Riil

Variabel	Rata-rata	Std. dev.	Minimum	Maksimum	Range
y_1	12,12	5,47	1,33	27,87	26,54
y_2	73,20	3,63	61,67	80,24	18,57
x_1	5,98	1,11	0,10	8,98	8,88
x_2	48,48	5,60	35,60	63,16	27,56
x_3	57,35	4,65	44,84	65,27	20,43
x_4	6,36	2,94	1,02	15,92	14,90
x_5	27,63	20,15	0,11	79,16	79,16

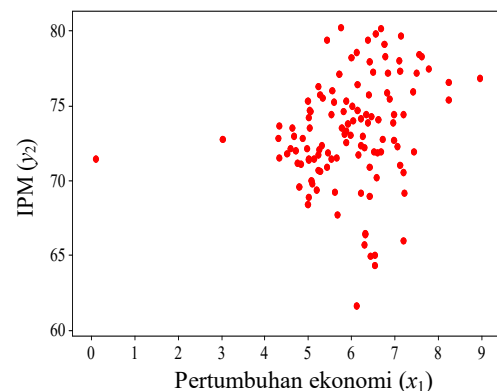
Berdasarkan informasi Tabel 6.1 pada kolom minimum dan maksimum, maka nilai persentase penduduk miskin (y_1) terendah berada pada Kota Tangerang Selatan sedangkan persentase tertinggi berada pada Kabupaten Sampang. Nilai IPM (y_2) terendah berada pada Kabupaten Sampang sedangkan tertinggi berada pada Kota Yogyakarta. Nilai pertumbuhan ekonomi (x_1) terendah berada pada Kabupaten Kepulauan Seribu sedangkan tertinggi berada pada Kota Bandung. Angka beban ketergantungan (x_2) terendah berada pada Kota Yogyakarta sedangkan tertinggi berada pada Kabupaten Garut. Nilai persentase pengeluaran perkapita untuk makanan (x_3) terendah berada pada Kota Jakarta Selatan sedangkan persentase tertinggi pada Kabupaten Bangkalan. Nilai TPT (x_4) terendah terletak di Kabupaten Pacitan sedangkan tertinggi di Kota Cirebon. Persentase penduduk yang perkerjaan utamanya di sektor pertanian (x_5) terendah berada pada Kota Jakarta Selatan sedangkan tertinggi di Kabupaten Sumenep.

Selain nilai minimum dan maksimum data penelitian yang disajikan pada tabel 6.1, terdapat pula beberapa data sangat ekstrim (*outlier*). Beberapa diantaranya adalah data pertumbuhan ekonomi di Kabupaten Kepulauan seribu

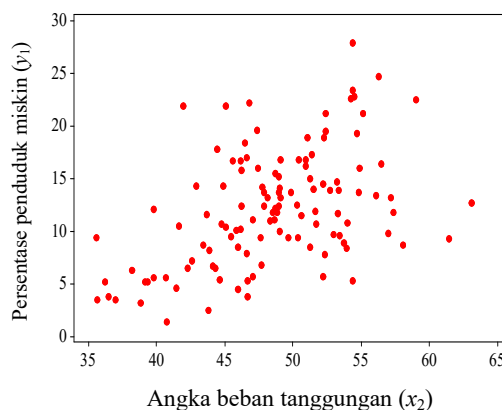
(0,1%) dan Kabupaten Cilacap (3,03%). Rendahnya pertumbuhan ekonomi disebabkan karena sumber PDRB (komponen yang digunakan untuk menghitung pertumbuhan ekonomi) di daerah ini lebih didominasi oleh pendapatan dari sektor minyak dan gas, daripada sektor-sektor lainnya (perdagangan, jasa dan lain-lain). Masing-masing untuk kepulauan Seribu sebesar 83,93% dan Cilacap 42,95%. Hanya saja pertumbuhan ekonomi sektor minyak dan gas di daerah ini sangat rendah dibandingkan dengan pertumbuhan ekonomi di sektor lainnya, masing-masing sebesar -0,89% untuk Kepulauan Seribu dan -0,18% untuk Kabupaten Cilacap. Akibatnya secara agregat pertumbuhan ekonomi di daerah tersebut menjadi sangat rendah. Selengkapnya, pola hubungan antara respon dan prediktor disajikan secara *scatter plot* dalam Gambar 6.1. Dari gambar ini, dapat diidentifikasi secara awal bagaimana bentuk pola hubungan dari variabel persentase penduduk miskin dan IPM dengan variabel-variabel yang mempengaruhinya.



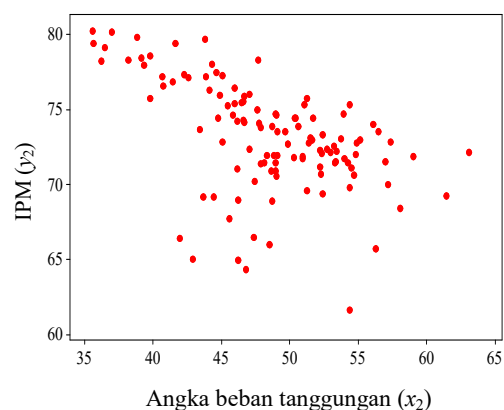
(a) *scatter plot* y_1 dengan x_1



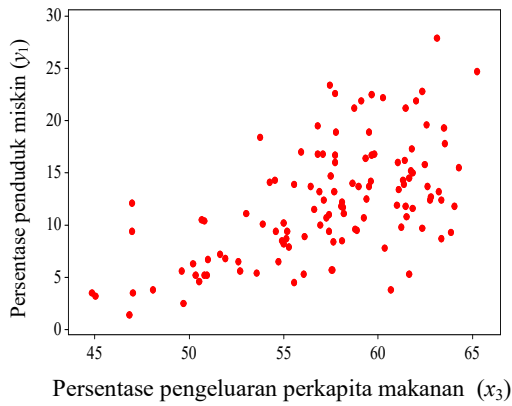
(f) *scatter plot* y_2 dengan x_1



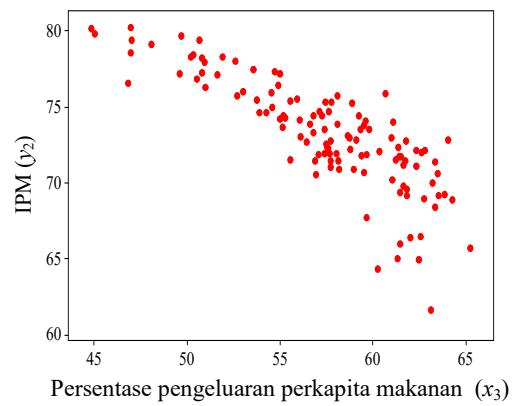
(b) *scatter plot* y_1 dengan x_2



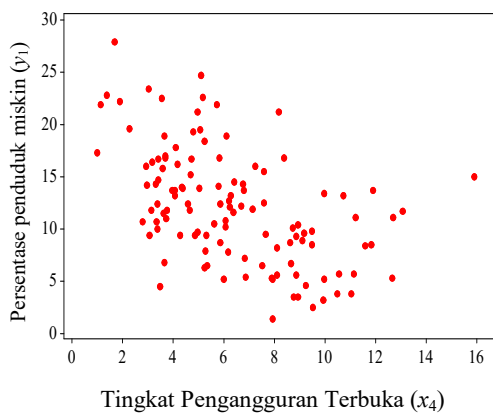
(g) *scatter plot* y_2 dengan x_2



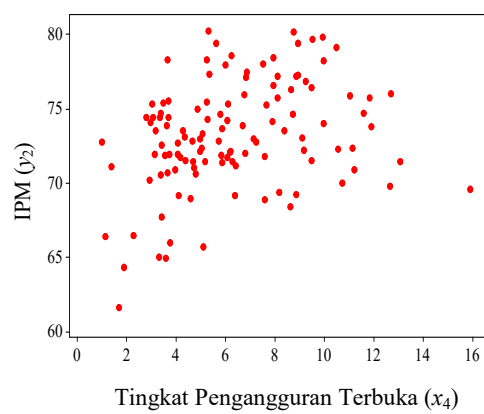
(c) *scatter plot* y_1 dengan x_3



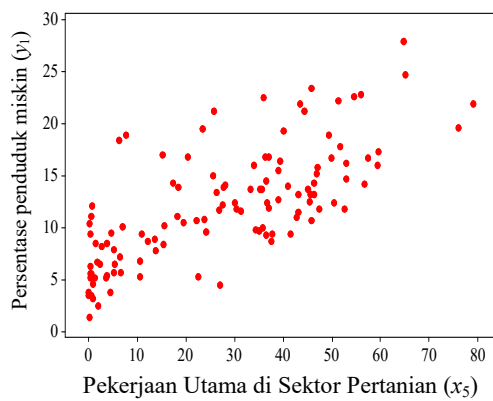
(h) *scatter plot* y_2 dengan x_3



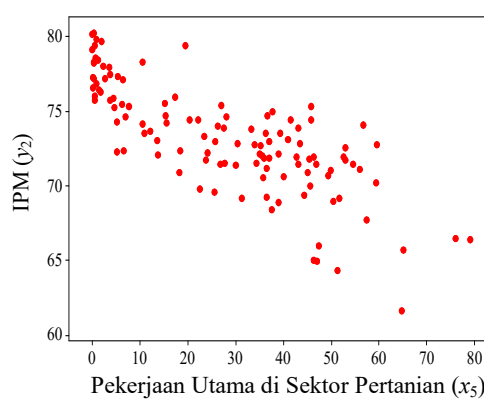
(d) *scatter plot* y_1 dengan x_4



(i) *scatter plot* y_2 dengan x_4



(e) *scatter plot* y_1 dengan x_5



(j) *scatter plot* y_2 dengan x_5

Gambar 6.1 *Scatter Plot* Respon dengan Prediktor pada Data Pemodelan
Persentase Penduduk Miskin dan IPM di Pulau Jawa

Hasil plot antara masing-masing prediktor terhadap respon yang ditunjukkan pada Gambar 6.1, mengindikasikan bahwa terdapat beberapa plot yang tidak menunjukkan pola yang jelas atau ada kecenderungan plot data

memiliki perubahan perilaku pada suatu interval tertentu. Sebagai contoh pada plot hubungan antara variabel TPT (x_4) terhadap variabel persentase penduduk miskin (y_1). Banyak literatur yang menyatakan bahwa terdapat korelasi antara TPT dan kemiskinan yang menyatakan bahwa TPT berpengaruh positif terhadap kemiskinan. Secara teori jika masyarakat tidak menganggur berarti mempunyai pekerjaan dan penghasilan, dan dengan penghasilan yang dimiliki dari bekerja diharapkan dapat memenuhi kebutuhan hidup. Jika kebutuhan hidup terpenuhi, maka tidak akan miskin. Sehingga dikatakan dengan tingkat pengangguran rendah (kesempatan kerja tinggi) maka tingkat kemiskinan juga rendah.

Namun jika diamati secara seksama, teori tersebut tidak selalu sesuai dengan fenomena yang terjadi di lapangan. Menurut *scatter plot* dalam Gambar 6.1 (d) dan data Lampiran 15, tampak bahwa terdapat perubahan perilaku data pada sub interval tertentu dalam kurva regresi tersebut. Sebagai contoh pada interval $x_4 < 11$, TPT (x_4) memberikan korelasi negatif terhadap persentase penduduk miskin (y_1). Hal ini dapat dijelaskan dengan fenomena sebagai berikut: kabupaten dengan TPT terendah biasanya merupakan daerah yang sebagian besar tenaga kerjanya bekerja pada sektor pertanian yang melibatkan hampir seluruh anggota keluarga (sehingga tingkat pengangguran rendah) tetapi dengan penghasilan yang rendah sehingga tidak mencukupi kebutuhan keluarga. Sehingga walaupun dengan TPT yang rendah (karena sebagian besar bekerja), namun mereka tetap miskin. Sementara untuk daerah perkotaan sering ditemui kondisi dimana tingkat pengangguran meningkat tetapi tingkat kemiskinan justru tidak ikut meningkat. Hal ini disebabkan karena meskipun tingkat pengangguran tinggi tetapi ada anggota rumah tangga lain yang bekerja dengan tingkat pendapatan tinggi sehingga cukup untuk menyokong penganggur.

Fenomena ini tidak terjadi pada semua kabupaten/kota. *Scatter plot* pada Gambar 6.1 (d) dan data lampiran 15 menunjukkan perubahan perilaku pada interval $x_4 > 11$, dimana TPT tidak lagi berpengaruh negatif tetapi cenderung berpengaruh positif terhadap kemiskinan. Kondisi ini sesuai dengan teori yang telah disebutkan sebelumnya bahwa kenaikan TPT memberikan pengaruh terhadap peningkatan tingkat kemiskinan. Jadi, untuk kasus kabupaten/kota di

Pulau Jawa, hubungan searah antara TPT dan kemiskinan di wilayah ini hanya terjadi sebagian, sementara pada beberapa kabupaten/kota lainnya tidak searah.

Kondisi yang sama terjadi pula pada pola hubungan antara variabel prediktor TPT (x_4) terhadap variabel IPM (y_2) yang ditunjukkan pada Gambar 6.1 (i). Secara teori jika masyarakat tidak menganggur berarti mempunyai pekerjaan dan penghasilan, dan dengan penghasilan yang dimiliki dari bekerja diharapkan dapat memenuhi kebutuhan hidup berupa pendidikan, kesehatan dan kehidupan layak yang diukur dengan IPM. Sehingga dikatakan dengan tingkat pengangguran rendah (kesempatan kerja tinggi) maka angka IPM akan tinggi. Namun fenomena yang terjadi di daerah tidak selalu sesuai dengan teori yang disebutkan diatas. *Scatter plot* pada Gambar 6.1 (i) dan data lampiran 15 mengindikasikan bahwa terdapat perubahan pola perilaku data pada subinterval tertentu dalam kurva regresi tersebut. Sebagai contoh pada interval $x_4 < 11$, angka TPT (x_4) cenderung berkorelasi positif terhadap peningkatan angka IPM (y_2).

Hal ini dapat dijelaskan dengan fenomena yang sama seperti pada hubungan antara TPT dengan kemiskinan. Kabupaten/kota dengan tingkat TPT tinggi biasanya adalah daerah pertanian yang sebagian besar tenaga kerjanya melibatkan anggota rumah tangganya sehingga pengangguran berkurang, namun dengan penghasilan yang rendah yang tidak dapat mencukupi kebutuhan keluarga. Akibatnya kebutuhan akan pendidikan, kesehatan dan kehidupan yang layak yang merupakan komponen dari angka IPM tidak terpenuhi. Sebaliknya pada daerah perkotaan, meskipun TPT nya tinggi, tetapi kebutuhan akan pendidikan, kesehatan dan kehidupan layak tetap dapat dipenuhi karena ada anggota keluarga lainnya yang dapat menyokong kebutuhan tersebut. Namun fenomena ini tidak terjadi pada semua kabupaten/kota. *Scatter plot* pada Gambar 6.1 (i) dan data lampiran 15 menunjukkan bahwa pada interval $x_4 > 11$, TPT cenderung berpengaruh negatif terhadap IPM. Kondisi ini sesuai dengan teori yang telah disebutkan sebelumnya bahwa kenaikan TPT memberikan pengaruh terhadap penurunan angka IPM di suatu wilayah. Hal ini berarti untuk kasus di Pulau Jawa, hubungan searah antara TPT dan IPM di wilayah ini hanya terjadi sebagian, sementara pada beberapa kabupaten/kota lainnya hubungannya tidak searah.

Sementara itu untuk pola hubungan pada variabel lainnya, yaitu pada pola angka pertumbuhan ekonomi (x_1), angka beban ketergantungan (x_2), angka persentase pengeluaran perkapita makanan (x_3), angka persentase penduduk yang pekerjaan utamanya di sektor pertanian (x_5) terhadap angka persentase penduduk miskin (y_1) dan angka IPM (y_2) yang disajikan dalam Gambar 6.1 umumnya telah sesuai dengan teori yang ada. Masing-masing pola tersebut menunjukkan bahwa: (1) peningkatan angka pertumbuhan ekonomi memberikan pengaruh pada penurunan angka persentase penduduk miskin dan peningkatan angka IPM di suatu wilayah, (2) penurunan angka beban ketergantungan memberikan pengaruh pada penurunan angka persentase penduduk miskin dan peningkatan angka IPM di suatu wilayah, (3) penurunan angka persentase pengeluaran perkapita makanan memberikan pengaruh pada penurunan angka persentase penduduk miskin dan peningkatan angka IPM di suatu wilayah dan (4) peningkatan angka persentase penduduk yang pekerjaan utamanya di sektor pertanian memberi pengaruh pada peningkatan persentase penduduk miskin dan penurunan IPM di suatu wilayah.

Meskipun pola hubungan antara variabel x_1, x_2, x_3, x_5 terhadap y_1, y_2 sudah sesuai dengan teori yang ada, namun pola tersebut tidak sama untuk setiap interval nilai yang ada dalam setiap variabel tersebut. Sebagai contoh pada Gambar 6.1 (a) dan 6.1 (f) dalam interval $x_1 > 6$, pola angka pertumbuhan ekonomi terhadap persentase penduduk miskin menunjukkan penurunan, sedangkan terhadap angka IPM menunjukkan kenaikan, tetapi pada interval $x_1 < 6$ pola tersebut menyebar. Gambar 6.1 (b) dan 6.1 (g) dalam interval $x_2 < 50$, pola angka beban tanggungan terhadap persentase penduduk miskin menunjukkan kenaikan, sedangkan terhadap angka IPM menunjukkan penurunan, tetapi pada interval $x_2 > 50$ pola tersebut menyebar. Gambar 6.1 (c) dalam interval $x_3 < 58$, pola angka persentase pengeluaran perkapita makanan terhadap angka persentase penduduk miskin menunjukkan kenaikan, tetapi pada interval $x_3 > 58$ pola tersebut menyebar. Gambar 6.1 (h) dalam interval $x_3 > 55$, menunjukkan bahwa pola angka persentase pengeluaran perkapita makanan terhadap angka IPM menunjukkan penurunan yang lebih cepat dibandingkan pada

interval $x_3 < 55$. Selain itu pada Gambar 6.1 (e) dan 6.1 (j) dalam interval $x_5 < 5$, menunjukkan bahwa pola angka persentase penduduk yang pekerjaannya utamanya di sektor pertanian menunjukkan kenaikan yang lebih cepat terhadap peningkatan angka persentase penduduk miskin dan penurunan yang lebih cepat terhadap penurunan angka IPM, dibandingkan pada pola interval $x_5 > 5$.

Adanya keterbatasan informasi mengenai bentuk pola hubungan dan bentuk perubahan perilaku kurva regresi dalam beberapa variabel penelitian yang disebutkan diatas, merupakan pertimbangan digunakannya pendekatan regresi nonparametrik untuk memodelkan pola data. Selain itu dari hasil uji linieritas *Ramsey Reset Test* dengan toleransi $p\text{-value} > 0,10$ yang disajikan pada Tabel 6.2 memperlihatkan bahwa hubungan antara prediktor terhadap respon adalah hubungan tidak linier. Hal ini memberikan alasan untuk menggunakan analisis regresi nonparametrik dalam penelitian, salah satunya dengan metode MARS.

Tabel 6.2. Uji Non-Linieritas *Ramsey Reset Test* pada Data Riil

Hubungan	F	$p\text{-value}$	Kesimpulan
x_1, x_2, x_3, x_4, x_5 terhadap y_1	3,8201	0,0249	Tidak linier
x_1, x_2, x_3, x_4, x_5 terhadap y_2	3,1585	0,0783	Tidak linier

Selanjutnya dari hasil pengujian korelasi *pearson* antar respon yaitu angka persentase penduduk miskin dan angka IPM menunjukkan adanya korelasi negatif yang signifikan yaitu $-0,6242$ dengan $p\text{-value}$ sebesar 0.000, sehingga pemodelan regresi MARS birespon kontinu merupakan salah satu solusi yang tepat dalam memodelkan pola data.

6.2. Model Persentase Penduduk Miskin dan IPM

Tahapan pemodelan MARS birespon kontinu diawali dengan mencari fungsi basis optimal dengan proses *stepwise* (*forward* dan *backward*) berdasarkan nilai GCV minimum pada masing-masing respon. Untuk mengakomodir adanya korelasi antar respon, maka diestimasi pula matriks bobot berdasarkan fungsi basis optimal yang diperoleh dari masing-masing respon menggunakan persamaan (4.25). Selanjutnya dari estimasi matriks bobot dan fungsi basis optimal masing-

masing respon tersebut, diperoleh estimasi parameter model MARS birespon kontinu menggunakan persamaan (4.23). Berdasarkan tahapan diatas maka dapat diperoleh fungsi basis optimal dan estimasi parameter MARS birespon kontinu untuk pemodelan persentase penduduk miskin dan IPM, yang hasilnya disajikan dalam Tabel 6.3.

Tabel 6.3. Estimasi Parameter dan Fungsi Basis Optimal pada Pemodelan Persentase Penduduk Miskin dan IPM di Pulau Jawa

Respon	Fungsi Basis	Nilai Fungsi Basis	Parameter
Respon 1	BF ₁₀	1 (intersep)	8,18
	BF ₁₁	$(x_{4i} - 11,14)_+$	1,65
	BF ₁₂	$(11,14 - x_{4i})_+$	0,90
	BF ₁₃	$(x_{1i} - 5,54)_+ (6 - x_{4i})_+$	-1,73
	BF ₁₄	$(x_{1i} - 5,54)_+ (6 - x_{4i})_+ (x_{2i} - 53,28)_+$	3,14
	BF ₁₅	$(x_{1i} - 5,54)_+ (2,98 - x_{4i})_+$	2,81
	BF ₁₆	$(59,68 - x_{3i})_+$	-0,46
	BF ₁₇	$(x_{3i} - 59,68)_+ (46,31 - x_{5i})_+$	-0,08
	BF ₁₈	$(x_{5i} - 4,73)_+$	0,10
Respon 2	BF ₂₀	1 (intersep)	73,31
	BF ₂₁	$(x_{3i} - 50,19)_+ (x_{5i} - 6,42)_+$	-0,01
	BF ₂₂	$(x_{3i} - 50,19)_+ (6,42 - x_{5i})_+$	0,12
	BF ₂₃	$(x_{3i} - 50,19)_+ (x_{5i} - 6,42)_+ (x_{1i} - 4,81)_+$	-0,002
	BF ₂₄	$(11,14 - x_{4i})_+$	-0,30
	BF ₂₅	$(55,28 - x_{3i})_+$	0,62
	BF ₂₆	$(11,14 - x_{4i})_+ (x_{3i} - 54,28)_+ (51,52 - x_{2i})_+$	-0,02
	BF ₂₇	$(x_{3i} - 50,19)_+ (x_{5i} - 6,42)_+ (54,04 - x_{2i})_+$	0,001

Berdasarkan estimasi parameter dan fungsi basis optimal dalam Tabel 6.3, maka dapat dituliskan persamaan regresi MARS birespon kontinu untuk pemodelan persentase penduduk miskin dan IPM sebagai berikut :

$$\begin{aligned} \hat{f}_1(x_{1i}, \dots, x_{5i}) = & 8,18 + 1,65(x_{4i} - 11,14)_+ + 0,90(11,14 - x_{4i})_+ - (1,73(x_{1i} - 5,54)_+ \times \\ & (6,00 - x_{4i})_+) + 3,14(x_{1i} - 5,54)_+ (6,00 - x_{4i})_+ (x_{2i} - 53,28)_+ + (2,81 \times \\ & (x_{1i} - 5,54)_+ (2,98 - x_{4i})_+) - 0,46(59,68 - x_{3i})_+ - (0,08(x_{3i} - 59,68)_+ \times \\ & (46,31 - x_{5i})_+) + 0,10(x_{5i} - 4,73)_+, \end{aligned}$$

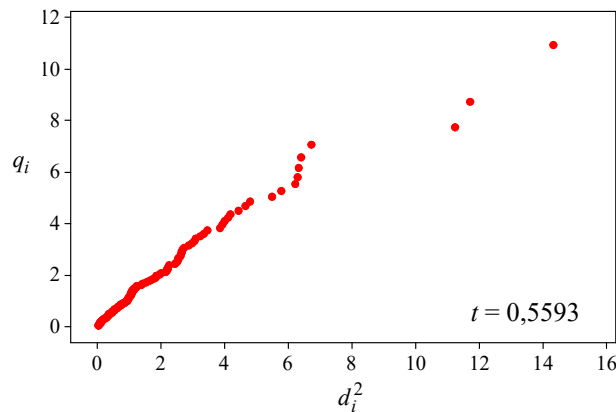
$$\begin{aligned}\hat{f}_2(x_{1i}, \dots, x_{5i}) = & 73,31 - 0,01(x_{3i} - 50,19)_+(x_{5i} - 6,42)_+ + 0,12(x_{3i} - 50,19)_+(6,42 - x_{5i})_+ - \\ & 0,002(x_{3i} - 50,19)_+(x_{5i} - 6,42)_+(x_{1i} - 4,81)_+ - 0,30(11,14 - x_{4i})_+ + (0,62 \times \\ & (55,28 - x_{3i})_+) - 0,02(11,14 - x_{4i})_+(x_{3i} - 54,28)_+(51,52 - x_{2i})_+ + (0,001 \times \\ & (x_{3i} - 50,19)_+(x_{5i} - 6,42)_+(54,04 - x_{2i})_+).\end{aligned}\quad (6.1)$$

Untuk menguji residual berdistribusi normal multivariat seperti yang di asumsikan dalam pemodelan ini, maka dilakukan uji Q-Q Plot dengan hipotesis sebagai berikut:

H_0 : Residual berdistribusi normal multivariat

H_1 : Residual tidak berdistribusi normal multivariat

Daerah penolakan apabila $t < 0.5$, penjelasan mengenai nilai t di berikan dalam Bab 2.

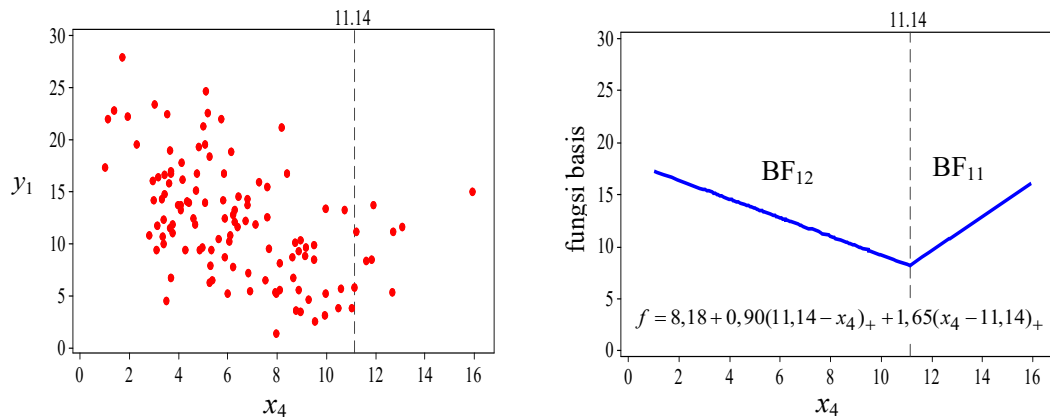


Gambar 6.2 Plot Uji Q-Q Plot pada Residual Model

Gambar 6.2 menunjukkan bahwa plot dari uji Q-Q Plot pada residual model cenderung membentuk garis lurus, selain itu ada lebih dari 50% ($t = 55,93\%$) nilai $d_i^2 \leq \chi_{2;0,50}^2$, sehingga kesimpulannya gagal tolak H_0 atau dapat dinyatakan bahwa residual model cenderung berdistribusi normal multivariat.

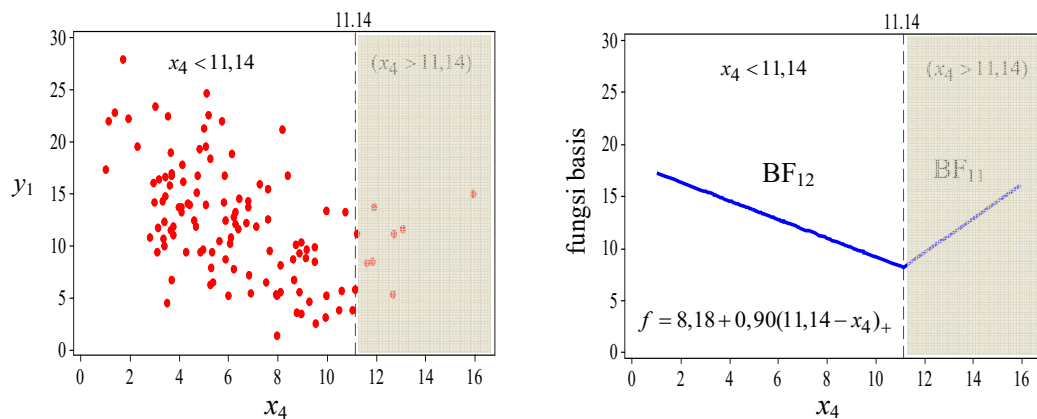
Contoh penjelasan fungsi basis dalam pemodelan persentase penduduk miskin dan IPM pada persamaan (6.1), diberikan oleh fungsi basis BF_{11} dan BF_{12} , yang menunjukkan pola hubungan antara tingkat pengangguran (x_4) dengan persentase penduduk miskin (y_1), tanpa melibatkan interaksi dengan variabel x

(prediktor) lainnya. Bentuk pola hubungan berdasarkan fungsi basis tersebut disajikan dalam Gambar 6.3.



Gambar 6.3.. Bentuk Fungsi Basis BF_{11} dan BF_{12}

Pembahasan kedua contoh fungsi basis dalam Gambar 6.3 diatas, akan dimulai dengan penjelasan pola hubungan fungsi basis BF_{12} (Gambar 6.4) dilanjutkan dengan penjelasan fungsi basis BF_{11} (Gambar 6.5) dan kemudian akan dijelaskan pula alasan terjadinya perbedaan pola fungsi basis tersebut dalam hubungan antara persentase penduduk miskin dan tingkat pengangguran terbuka.



Gambar 6.4. Bentuk Fungsi Basis BF_{12}

Fungsi basis BF_{12} pada Gambar 6.4, menyatakan bahwa pada segmen/wilayah $x_4 < 11.14$, setiap perubahan 1% tingkat pengangguran terbuka (x_4) akan meningkatkan persentase penduduk miskin sebesar 0,90%. Hal ini berarti bahwa semakin rendah tingkat pengangguran terbuka (TPT), maka

semakin tinggi persentase penduduk miskin. Pola ini menunjukkan bahwa hubungan antara pengangguran dan kemiskinan tidak selalu sesuai dengan asumsi teori yang ada, bahwa semakin rendah pengangguran maka semakin rendah tingkat kemiskinan, tetapi memiliki hubungan yang terbalik. Terdapat beberapa penelitian di Indonesia maupun di luar negeri yang dapat dijadikan acuan untuk menjelaskan pola hubungan ini.

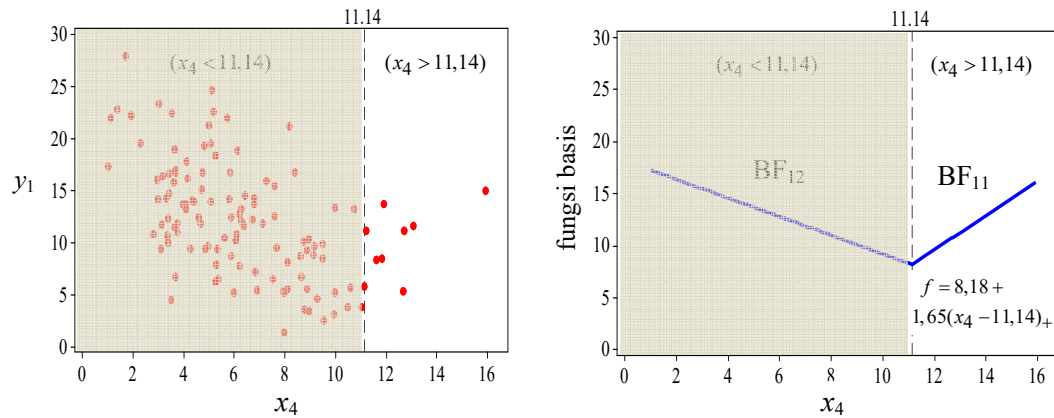
- a. Ratnasari (2014) dalam penelitiannya tentang pengaruh tingkat pengangguran terhadap tingkat kemiskinan di Indonesia, menyatakan bahwa pada kelompok keluarga miskin, sebagian besar anggota keluarga bekerja untuk bisa bertahan hidup, terkadang anak-anak juga dilibatkan dalam bekerja dengan alasan penghasilan kepala keluarga atau orang tua tidak mencukupi kebutuhan keluarga, sehingga adanya kecenderungan tingkat pengangguran akan rendah. Gita (2014) juga menyatakan bahwa di daerah perkotaan, seringkali ditemui fakta bahwa tingkat pengangguran tinggi tetapi kemiskinan tidak ikut meningkat. Hal ini karena kadangkala ada penduduk di perkotaan yang menganggur secara sukarela karena mencari pekerjaan yang lebih baik yang lebih sesuai dengan tingkat pendidikannya. Mereka menolak pekerjaan yang mereka rasakan lebih rendah dan mereka bersikap demikian karena mempunyai sumber lain yang bisa membantu masalah keuangan mereka. Fakta lainnya, seringkali ditemui anggota keluarga yang baru saja menamatkan pendidikan, tetapi tidak serta merta langsung bekerja karena masa tunggu mendapatkan pekerjaan relatif lama. Dalam kaitannya dengan kemiskinan, penganggur yang ada di rumah tangga tersebut tidak secara otomatis menjadi miskin karena ada anggota keluarga lain yang memiliki pendapatan yang cukup untuk mempertahankan keluarganya hidup berada diatas garis kemiskinan. Tidak hanya itu, ada banyak kegiatan penduduk yang menurut survei tidak di klaim sebagai bekerja (menganggur), tetapi justru pendapatan mereka umumnya diatas garis kemiskinan, seperti pekerja seks komersil, penjual barang ilegal (narkoba), perampok, pengemis, dan lain-lain. Hal-hal tersebut juga menjadi penyebab tingkat pengangguran meningkat tetapi tingkat kemiskinan berkurang.

- b. Penjelasan lain oleh Oshima, 1990 menyatakan bahwa rumahtangga miskin hampir tidak mungkin menjadi penganggur. Pernyataan Oshima tersebut dapat dipahami mengingat di negara berkembang seperti Indonesia tidak terdapat jaminan sosial bagi penganggur, sehingga orang miskin untuk bertahan hidup mau tidak mau harus bekerja meskipun hanya beberapa jam seminggu.
- c. Penelitian lain yaitu Yacoub (2012), yang mengkaji pengaruh tingkat pengangguran terhadap tingkat kemiskinan kabupaten/kota di Propinsi Kalimantan Barat. Hasil penelitian menyatakan bahwa kabupaten dengan TPT terendah biasanya merupakan daerah yang sebagian besar tenaga kerjanya bekerja pada sektor pertanian yang melibatkan hampir seluruh anggota keluarga (sehingga tingkat pengangguran rendah) tetapi dengan penghasilan yang rendah sehingga tidak mencukupi kebutuhan keluarga. Sehingga walaupun dengan TPT yang rendah (karena sebagian besar bekerja), namun mereka tetap miskin.
- d. Penelitian lainnya dilakukan oleh Osinubi (2005), yang menyatakan bahwa terdapat korelasi tinggi antara pengangguran dan persentase penduduk miskin di Nigeria. Hasil penelitian Osinubi (2005) menyatakan bahwa penurunan pengangguran meningkatkan penduduk miskin dan sebaliknya peningkatan pengangguran mengurangi kemiskinan di Nigeria. Hal ini disebabkan karena penduduk yang sangat miskin cenderung menerima pekerjaan apa saja, meskipun dengan gaji rendah, sehingga pengangguran berkurang. Sedangkan untuk penduduk yang menganggur cenderung ditanggung oleh keluarga mereka yang tidak miskin, sehingga tidak secara otomatis menjadi miskin.
- e. Streeten *et al.* (1981), yang melakukan penelitian kebutuhan dasar penduduk di negara berkembang, secara eksplisit juga menyatakan bahwa orang yang menganggur belum tentu miskin dan orang yang bekerja belum tentu kaya. Karena umumnya orang yang menganggur memiliki kondisi yang lebih baik, sementara kebanyakan orang yang benar-benar miskin tidak menganggur.

Berdasarkan uraian para peneliti diatas dan dari pola fungsi basis BF₁₂ (Gambar 6.4), menyatakan bahwa semakin rendah pengangguran maka tingkat kemiskinan cenderung semakin tinggi. Contoh kabupaten menurut pola tersebut,

diantaranya adalah Kabupaten Sumenep dengan tingkat pengangguran (x_4) 1,14% dan penduduk miskin 21,87%. Kabupaten Gunung Kidul dengan tingkat pengangguran 1,38% dan penduduk miskin 22,72%. Kabupaten Sampang dengan tingkat pengangguran 1,71% dan penduduk miskin 27,87%. Kabupaten Probolinggo dengan tingkat pengangguran 1,92% dan penduduk miskin 22,14%. Menurut para peneliti banyak faktor yang dapat menyebabkan hal tersebut, salah satunya adalah tingginya persentase penduduk pertanian di daerah itu (Yacoub, 2012). Masing-masing Kabupaten Sumenep dengan penduduk pertanian 79,16%, Kabupaten Gunung Kidul dengan penduduk pertanian 55,96%, Kabupaten Sampang dengan penduduk pertanian 64,79% dan Kabupaten Probolinggo dengan penduduk pertanian 51,41%.

Pola pada fungsi basis BF_{12} juga menunjukkan bahwa semakin tinggi pengangguran maka tingkat kemiskinan cenderung semakin rendah. Contoh kabupaten yang mewakili pola tersebut, diantaranya adalah Jakarta Selatan dengan tingkat pengangguran 8,77% dan penduduk miskin 3,49%. Jakarta Timur dengan tingkat pengangguran 9,95% dan penduduk miskin 3,12%. Jakarta Pusat dengan tingkat pengangguran 10,5% dan penduduk miskin 3,72%. Jakarta Barat dengan tingkat pengangguran 8,95% dan penduduk miskin 3,47%. Jakarta Utara dengan tingkat pengangguran 9,98% dan penduduk miskin 5,14%. Depok dengan tingkat pengangguran 9,55% dan penduduk miskin 2,46%. Tangerang dengan tingkat pengangguran 8,13% dan penduduk miskin 5,55%. Bekasi dengan tingkat pengangguran 8,88% dan penduduk miskin 5,55%. Menurut Yacoub (2012) biasanya pekerjaan utama penduduk di daerah tersebut bukan dari sektor pertanian, masing-masing Jakarta Selatan dengan penduduk pertanian 0,11%, Jakarta Timur dengan penduduk pertanian 0,96%, Jakarta Pusat dengan penduduk pertanian 0%, Jakarta Barat dengan penduduk pertanian 0,62%, Jakarta Utara dengan penduduk pertanian 0,40%, Depok dengan penduduk pertanian 2,02%, Tangerang dengan penduduk pertanian 0,67% dan Bekasi dengan penduduk pertanian 0,41%. Meskipun demikian, pola tersebut tidak semata-mata hanya dipengaruhi oleh rendahnya penduduk pertanian di daerah tersebut, tetapi disebabkan oleh banyak faktor seperti yang diuraikan oleh para peneliti sebelumnya.



Gambar 6.5. Bentuk Fungsi Basis BF_{11}

Penjelasan berikutnya pada fungsi basis BF_{11} dalam Gambar 6.5, menyatakan bahwa pada interval $x_4 > 11.14$, setiap perubahan 1% tingkat pengangguran terbuka (x_4) akan meningkatkan persentase penduduk miskin sebesar 1,65%. Pola fungsi basis ini menunjukkan bahwa TPT memiliki hubungan positif terhadap persentase penduduk miskin. Berdasarkan scatter plot, tampak beberapa kabupaten yang mengikuti pola ini diantaranya adalah Kabupaten Bandung dengan tingkat pengangguran 11,61% dan penduduk miskin 8,32%, Kota Sukabumi dengan tingkat pengangguran 11,85% dan penduduk miskin 8,41%, Kota Cirebon dengan tingkat pengangguran 12,71% dan penduduk miskin 11,08%, Kepulauan Seribu dengan tingkat pengangguran 13,08% dan penduduk miskin 11,62% serta Kabupaten Cirebon dengan tingkat pengangguran 15,92% dan penduduk miskin 14,94%. Biasanya persentase penduduk dengan pekerjaan utama pertanian di daerah ini rendah. Masing-masing persentase penduduk miskin Kabupaten Bandung 15,53%, Kota Sukabumi 3,85%, Kota Cirebon 0,69%, Kepulauan Seribu 26,84% dan Kabupaten Cirebon 25,63%.

Selain itu tingginya tingkat pengangguran di daerah ini bisa jadi disebabkan karena terbatasnya lapangan kerja yang tersedia, yang dapat diukur dari rendahnya angka tingkat kesempatan kerja (TKK) di daerah ini. Masing-masing angka TKK Kabupaten Bandung 88,39%, Kota Sukabumi 88,15%, Kota Cirebon 87,29%, Kepulauan Seribu 86,92% dan Kabupaten Cirebon 84,08%. Angka TKK di kabupaten-kabupaten ini jauh dibawah angka rata-rata TKK di pulau Jawa yaitu sebesar 93,64%. Pola hubungan positif dari fungsi basis ini sama dengan pola hubungan TPT dan kemiskinan dalam penelitian Son dan Kakwani

(2006) di Thailand dan Korea. Pendapatan penduduk bekerja dalam penelitian Son dan Kakwani (2006) diatas upah minimum yang ditetapkan oleh negara tersebut, akibatnya semakin banyak penduk yang bekerja (semakin sedikit pengangguran), semakin sedikit penduduk miskin.

Adanya perbedaan pola hubungan antara respon dan prediktor dalam prediktor yang sama, seperti yang ditunjukkan oleh fungsi basis BF_{11} dan BF_{12} , seringkali akan ditemui dalam hasil pemodelan MARS. Hal ini disebabkan karena dalam MARS terdapat titik knot, sehingga sensitif terhadap perubahan pola data. Kondisi ini seringkali menimbulkan kesulitan dalam interpretasi data, sehingga perlu berhati-hati dalam menginterpretasikan model yang diperoleh. Sesuai dengan konsepnya, bahwa MARS adalah salah satu metode dalam regresi nonparametrik yang mengestimasi pola hubungan antara respon dan prediktor yang tidak diketahui bentuknya, maka hasil estimasinya berpeluang untuk menghasilkan fungsi basis yang menunjukkan hubungan antara prediktor dan respon yang tidak sesuai dengan teorinya. Hal ini berbeda dengan pendekatan regresi parametrik, yang mengestimasi pola hubungan yang telah diketahui bentuknya, berdasarkan teori-teori penelitian yang telah ada sebelumnya, sehingga tidak susah dalam melakukan interpretasi.

Meskipun demikian MARS memiliki kelebihan, karena menghasilkan estimasi model yang lebih baik daripada regresi parametrik, hal ini karena MARS memiliki titik knot sehingga lebih fleksibel dalam mengestimasi pola data. Berdasarkan hasil pengolahan diperoleh nilai R^2 dan $RMSE$ untuk model ini masing-masing sebesar 76,86% dan 2,2221. Hasil perbandingan estimasi MARS birespon dan regresi parametrik birespon berdasarkan nilai MSE , $RMSE$ dan R^2 *adjusted* disajikan dalam Tabel 6.4.

Tabel 6.4. Perbandingan Ketepatan Estimator Model MARS Birespon Kontinu dengan Model Parametrik pada Pemodelan Data Riil

Ukuran Ketepatan Estimator	Regresi Parametrik			MARS
	Linier	Kuadratik	Kubik	
MSE	7,4945	6,9765	6,7445	4,9377
$RMSE$	2,7376	2,6413	2,5970	2,2221
R^2 <i>adjusted</i>	62,98%	63,91%	63,39%	75,28%

Tabel 6.4 menunjukkan bahwa estimasi model MARS birespon kontinu lebih baik daripada estimasi model regresi nonparametrik, hal ini dikarenakan nilai MSE , $RMSE$ yang lebih kecil dan nilai R^2 *adjusted* model yang lebih besar dibandingkan dengan model regresi parametrik. Tingginya keakuratan model MARS birespon kontinu dalam persamaan (6.1), diharapkan dapat menjadikan model ini sebagai metode alternatif untuk untuk mengestimasi angka persentase penduduk miskin dan IPM di Pulau Jawa, dibandingkan dengan regresi parametrik yang banyak digunakan selama ini. Sehingga diharapkan setiap kebijakan yang akan dilakukan pemerintah untuk perbaikan persentase penduduk miskin dan IPM dapat dievaluasi dari hasil estimasi model ini.

6.3. Validasi Model Persentase Penduduk Miskin dan IPM

Hal yang perlu diperhatikan dalam pemodelan data adalah bagaimana membentuk model yang sesuai/tepat menggambarkan fenomena riil yang kompleks bukan mendapatkan model yang sempurna. Salah satu cara untuk mengetahuinya adalah dengan melakukan validasi model untuk melihat ketepatan model dalam prediksi. Ketepatan prediksi dalam model MARS dapat dievaluasi dengan *Root Mean Square Error of Cross Validation (RMSECV)* (Xu *et al.*, 2004).

Nilai $RMSECV$ dihitung dengan cara mengeluarkan satu observasi dari sekumpulan data (n) dan membangun model dari $n-1$ data yang tersisa. Model yang diperoleh selanjutnya digunakan untuk memprediksi data yang dikeluarkan. Ketika semua data telah dikeluarkan satu persatu dan dibangun model berdasarkan data yang tersisa, maka nilai $RMSECV$ diperoleh dari rata-rata nilai $RMSEP$ dari keseluruhan model. Rumus $RMSEP$ diberikan oleh persamaan (2.36) dalam Bab Kajian Pustaka dan Dasar Teori.

Untuk mengetahui pengaruh korelasi antar respon dalam ketepatan prediksi pemodelan persentase penduduk miskin dan IPM, maka dalam penelitian ini dibandingkan nilai $RMSECV$ antara model MARS birespon kontinu yang melibatkan korelasi dengan model yang tidak melibatkan korelasi atau diestimasi secara sendiri-sendiri. Nilai $RMSECV$ dari kedua model tersebut disajikan dalam Tabel 6.5.

Tabel 6.5. Perbandingan Nilai *RMSECV* dari Model yang Melibatkan Korelasi dengan Tidak Melibatkan Korelasi pada Data Riil

Respon	Nilai <i>RMSECV</i>	
	Model dengan Korelasi	Model Tanpa Korelasi
Birespon	2,7066	2,7640

Tabel 6.5 menunjukkan bahwa nilai *RMSECV* baik untuk model yang melibatkan korelasi maupun tidak melibatkan korelasi memiliki nilai yang tidak jauh berbeda dengan nilai *RMSE* model MARS birespon yang disajikan dalam Tabel 6.4. Dengan demikian dapat dinyatakan bahwa model yang dihasilkan telah valid untuk menggambarkan fenomena riil yang ada. Namun jika diperhatikan lebih seksama, tampak bahwa model yang melibatkan korelasi memiliki nilai *RMSECV* yang lebih kecil daripada model yang tidak melibatkan korelasi atau diestimasi secara sendiri-sendiri. Hal ini memberikan kesimpulan bahwa pemodelan persentase penduduk miskin dan IPM yang melibatkan korelasi antar respon menghasilkan tingkat prediksi yang lebih baik dibandingkan dengan jika tidak melibatkan korelasi atau dimodelkan sendiri-sendiri. Tingginya keakuratan prediksi dari model ini, diharapkan dapat menjadi alat evaluasi bagi kebijakan yang akan dilakukan oleh pemerintah untuk perbaikan kesejahteraan rakyat. Misalnya dengan cara mengevaluasi apakah kebijakan pemerintah yang terkait dengan program pertumbuhan ekonomi, angka beban tanggungan, persentase pengeluaran perkapita untuk makanan, tingkat pengangguran terbuka dan persentase penduduk yang pekerjaan utamanya disektor pertanian dapat memberi manfaat bagi penurunan angka persentase penduduk miskin dan peningkatan angka IPM di daerah tersebut dengan menggunakan metode estimasi ini.

BAB 7

KESIMPULAN DAN KEBERLANJUTAN

Pada bab ini diuraikan kesimpulan dari hasil pembahasan disertasi yang disajikan pada Bab 4, Bab 5 dan Bab 6 yang meliputi kajian teori pengembangan model MARS birespon kontinu, kajian terapan pada data simulasi dan penerapan model MARS birespon kontinu pada data riil. Selain itu juga disajikan beberapa keberlanjutan yang berkaitan dengan hasil penelitian ini.

7.1. Kesimpulan

Berdasarkan hasil dan pembahasan dapat disimpulkan beberapa hal sebagai berikut:

1. Model MARS birespon kontinu adalah pengembangan metode nonparametrik yang memodelkan pola hubungan yang tidak diketahui antara dua respon kontinu yang berkorelasi dengan p prediktor menggunakan fungsi MARS. Model MARS birespon kontinu dapat dinyatakan dalam bentuk matriks:

$$\underline{y} = \mathbf{B}(\underline{x}, \underline{t})\underline{\alpha} + \varepsilon,$$

dimana:

$$\underline{y} = (\underline{y}'_1, \underline{y}'_2)' = (y_{11}, y_{12}, \dots, y_{1n}, y_{21}, y_{22}, \dots, y_{2n})',$$

$$\mathbf{B}(\underline{x}, \underline{t}) = \begin{pmatrix} \mathbf{B}_1(\underline{x}, \underline{t}) & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_2(\underline{x}, \underline{t}) \end{pmatrix},$$

dengan :

$$\mathbf{B}_\ell(\underline{x}, \underline{t}) = \begin{pmatrix} 1 \prod_{k_\ell=1}^{K_1} [s_{k_\ell 1}(x_{v(k_\ell, 1)1} - t_{k_\ell 1})]_+ & \cdots & \prod_{k_\ell=1}^{K_{M_\ell}} [s_{k_\ell M_\ell}(x_{v(k_\ell, M_\ell)1} - t_{k_\ell M_\ell})]_+ \\ 1 \prod_{k_\ell=1}^{K_1} [s_{k_\ell 1}(x_{v(k_\ell, 1)2} - t_{k_\ell 1})]_+ & \cdots & \prod_{k_\ell=1}^{K_{M_\ell}} [s_{k_\ell M_\ell}(x_{v(k_\ell, M_\ell)2} - t_{k_\ell M_\ell})]_+ \\ \vdots & & \vdots \\ 1 \prod_{k_\ell=1}^{K_1} [s_{k_\ell 1}(x_{v(k_\ell, 1)n} - t_{k_\ell 1})]_+ & \cdots & \prod_{k_\ell=1}^{K_{M_\ell}} [s_{k_\ell M_\ell}(x_{v(k_\ell, M_\ell)n} - t_{k_\ell M_\ell})]_+ \end{pmatrix},$$

$\ell = 1, 2,$

$$x_{v(k_\ell, m_\ell)} \in \{x_j\}_{j=1}^p, t_{k_\ell m_\ell} \in \{x_{v(k_\ell, m_\ell)i}\}_{i=1}^n, m_\ell = 1, 2, \dots, M_\ell, \ell = 1, 2,$$

$$\underline{\alpha} = (\alpha'_1, \alpha'_2)' = (\alpha_{11}, \alpha_{12}, \dots, \alpha_{1M_1}, \alpha_{21}, \alpha_{22}, \dots, \alpha_{2M_2})',$$

$$\underline{\varepsilon} = (\varepsilon'_1, \varepsilon'_2)' = (\varepsilon_{11}, \varepsilon_{12}, \dots, \varepsilon_{1n}, \varepsilon_{21}, \varepsilon_{22}, \dots, \varepsilon_{2n})'.$$

Berdasarkan metode MLE diperoleh estimasi fungsi regresi untuk model MARS birespon kontinu sebagai berikut:

$$\hat{f}_{\underline{\alpha}}(x) = \mathbf{B}(x, t)(\mathbf{B}'(x, t)\mathbf{W}^{-1}\mathbf{B}(x, t))^{-1}\mathbf{B}'(x, t)\mathbf{W}^{-1}\underline{y}.$$

2. Estimasi fungsi regresi MARS birespon kontinu memuat matriks \mathbf{W} yang berfungsi sebagai matriks bobot dalam mengakomodasi adanya korelasi antar respon. Pada umumnya matriks \mathbf{W} tidak diketahui sehingga diestimasi dari data. Berdasarkan metode MLE diperoleh estimasi matriks \mathbf{W} sebagai berikut:

$$\hat{\mathbf{W}} = \begin{pmatrix} \hat{\sigma}_{11} & | & \hat{\sigma}_{12} \\ \hline \hat{\sigma}_{12} & | & \hat{\sigma}_{22} \end{pmatrix} \otimes \mathbf{I},$$

dengan:

$$\hat{\sigma}_{11} = \frac{\underline{y}'_1(\mathbf{I} - \mathbf{H}_1(\cdot))'(\mathbf{I} - \mathbf{H}_1(\cdot))\underline{y}_1}{n},$$

$$\hat{\sigma}_{22} = \frac{\underline{y}'_2(\mathbf{I} - \mathbf{H}_2(\cdot))'(\mathbf{I} - \mathbf{H}_2(\cdot))\underline{y}_2}{n},$$

$$\hat{\sigma}_{12} = \frac{\underline{y}'_1(\mathbf{I} - \mathbf{H}_1(\cdot))'(\mathbf{I} - \mathbf{H}_2(\cdot))\underline{y}_2}{n},$$

$$\mathbf{H}_1(\cdot) = \mathbf{B}_1(x, t)(\mathbf{B}'_1(x, t)\mathbf{B}_1(x, t))^{-1}\mathbf{B}'_1(x, t),$$

$$\mathbf{H}_2(\cdot) = \mathbf{B}_2(x, t)(\mathbf{B}'_2(x, t)\mathbf{B}_2(x, t))^{-1}\mathbf{B}'_2(x, t).$$

3. Kajian terapan pada simulasi fungsi asli MARS menunjukkan bahwa:
 - a. Semakin besar ukuran sampel maka semakin dekat nilai estimasi parameter dan knot dengan nilai aslinya dengan penyebaran/varians hasil estimasi yang relatif lebih kecil (konsisten), sebaliknya semakin kecil

ukuran sampel maka semakin jauh nilai estimasi parameter dan knot dibandingkan dengan nilai aslinya.

- b. Semakin kecil varians *error* maka semakin dekat nilai estimasi parameter dan knot dengan nilai aslinya dengan penyebaran hasil estimasi yang relatif lebih kecil, sebaliknya semakin besar varians *error* maka semakin jauh nilai estimasi parameter dan knot dibandingkan dengan nilai aslinya.
 - c. Metode MARS birespon kontinu cenderung baik digunakan untuk ukuran sampel (n) diatas 50 observasi ($n \geq 50$), dan cenderung kurang baik digunakan untuk $n < 50$ dengan varians *error* yang besar.
 - d. Jika fungsi basis yang digunakan sangat sedikit, maka tingkat keakuratan model akan sangat rendah, yang ditandai oleh tingginya nilai *RMSE* dan rendahnya nilai R^2 model. Sebaliknya jika fungsi basis yang digunakan sangat banyak, maka model yang dihasilkan akan sangat kompleks. Akibatnya perlu dipilih fungsi basis yang optimal agar diperoleh estimator MARS birespon kontinu yang paling sesuai untuk data.
 - e. Semakin besar nilai mutlak korelasi antar respon maka semakin baik penggunaan estimator MARS birespon kontinu, yang ditandai oleh semakin rendahnya nilai *RMSEP* model yang diperoleh. Hasil simulasi menunjukkan bahwa nilai mutlak korelasi diatas 0,3 ($|\rho| > 0,3$) cenderung memberikan hasil prediksi yang lebih baik.
4. Kajian terapan pada simulasi berbagai fungsi percobaan menunjukkan bahwa:
- a. Estimator MARS birespon kontinu mampu digunakan untuk mengestimasi berbagai bentuk fungsi percobaan, seperti pada fungsi percobaan Trigonometri, Polinomial, dan Eksponensial, yang ditunjukkan melalui pendekatan grafik.
 - b. Jika fungsi basis yang digunakan sangat sedikit, maka tingkat keakuratan model akan sangat rendah, yang ditandai oleh tingginya nilai *RMSE* dan rendahnya nilai R^2 model. Sebaliknya jika fungsi basis yang digunakan sangat banyak, maka model yang dihasilkan akan sangat kompleks. Akibatnya perlu dipilih fungsi basis yang optimal agar diperoleh estimator MARS birespon kontinu yang paling sesuai untuk data.

- c. Semakin besar nilai mutlak korelasi antar respon maka semakin baik penggunaan estimator MARS birespon kontinu, yang ditandai oleh semakin rendahnya nilai *RMSEP* model yang diperoleh. Hasil simulasi menunjukkan bahwa nilai mutlak korelasi diatas 0,3 ($|\rho| > 0,3$) cenderung memberikan hasil prediksi yang lebih baik.

5. Kajian terapan pada data riil menunjukkan bahwa:

- a. Hasil estimasi MARS birespon kontinu pada pemodelan persentase penduduk miskin dan IPM memberikan tingkat akurasi yang cukup baik dengan nilai R^2 sebesar 76,86% dan nilai *RMSE* sebesar 2,2221. Hasil estimasi ini jauh lebih baik dibandingkan dengan hasil estimasi yang diperoleh dengan metode regresi parametrik (linier, kuadratik dan kubik) yang banyak digunakan selama ini oleh para praktisi untuk mengestimasi angka persentase penduduk miskin dan IPM.
- b. Hasil validasi model yang bertujuan untuk melihat ketepatan prediksi MARS birespon kontinu pada data riil, menunjukkan nilai *RMSECV* sebesar 2,7066. Nilai ini tidak jauh berbeda dengan nilai *RMSE* model sebesar 2,2221, yang berarti bahwa model yang dihasilkan telah valid untuk menggambarkan fenomena riil yang ada.
- c. Nilai *RMSECV* model MARS yang tidak melibatkan korelasi memiliki nilai yang lebih besar dibandingkan dengan nilai *RMSECV* model MARS birespon yang melibatkan korelasi, yaitu 2,7640 berbanding 2,7066 . Hal ini menunjukkan bahwa pemodelan persentase penduduk miskin dan IPM yang melibatkan korelasi antar respon menghasilkan tingkat prediksi yang lebih baik dibandingkan dengan pemodelan yang tidak melibatkan korelasi.
- d. Tingginya keakuratan estimasi model MARS birespon kontinu baik dibandingkan dengan regresi parametrik maupun dari metode MARS yang tidak melibatkan korelasi antar respon, diharapkan dapat menjadi alat evaluasi bagi kebijakan yang akan dilakukan oleh pemerintah untuk perbaikan kesejahteraan rakyat. Misalnya dengan cara mengevaluasi apakah kebijakan pemerintah yang terkait dengan program pertumbuhan

ekonomi, angka beban tanggungan, persentase pengeluaran perkapita untuk makanan, tingkat pengangguran terbuka dan persentase penduduk yang pekerjaan utamanya disektor pertanian dapat memberi manfaat bagi penurunan angka persentase penduduk miskin dan peningkatan angka IPM di daerah tersebut dengan menggunakan metode estimasi ini.

7.2. Keberlanjutan

Berdasarkan kesimpulan yang diperoleh dalam disertasi ini, terdapat beberapa permasalahan terbuka yang dapat dikembangkan sebagai tindak lanjut dari penelitian ini, diantaranya:

1. Model MARS birespon kontinu yang dikembangkan dalam penelitian ini masih terbatas pada data *cross section*, sehingga untuk penelitian berikutnya dapat dikembangkan pada data *time series* dan pada data longitudinal.
2. Penelitian dalam disertasi ini masih dibatasi pada respon kontinu, sehingga untuk penelitian berikutnya dapat dikembangkan untuk pemodelan MARS birespon kategori.
3. Penelitian dalam disertasi ini masih dibatasi untuk penggunaan dua respon (birespon), sehingga memungkinkan untuk pengembangan model yang melibatkan lebih dari dua respon (multirespon), baik untuk respon kontinu maupun pada respon kategori.
4. Kajian teori dalam disertasi ini masih terbatas pada bentuk estimator MARS birespon kontinu, sehingga perlu penelitian lanjutan untuk mengkaji inferensi statistika seperti penentuan statistik uji, sifat asimtotis dan interval konfidensi pada estimator MARS birespon kontinu.

Halaman ini sengaja dikosongkan

DAFTAR PUSTAKA

- Abraham, A. and Steinberg, D. (2001). *MARS: Still an Alien Planet in Soft Computing?*. School of Computing and Information Technology, Salford System. Inc, USA.
- Astuti, E.T., Budiantara, I.N., Sunaryo, S. and Dokhi, M. (2013). Statistical Modelling for Mortality Data using Local Generalized Poisson Regression Model. *International Journal of Applied Mathematics and Statistics*, 33(3), 92-101.
- Aydin, D. (2008). A Comparison of the Nonparametric Regression Models using Smoothing Spline and Kernel Regression. *Proc. World Acad. Sci. Eng. Technol.*, 36: 253-257.
- Ayomi, S. (2014). *Analisis Pengaruh Tingkat Pertumbuhan Sektor Utama terhadap Tingkat Kemiskinan, Tingkat Pengangguran dan Indeks Pembangunan Manusia (IPM) di Satuan Wilayah Madiun dan Sekitarnya Tahun 2003-2012*. Fakultas Ekonomi dan Bisnis, Universitas Brawijaya.
- Badan Pusat Statistik. (2013). *Data dan Informasi Kemiskinan Kabupaten/Kota Tahun 2012*. BPS, Jakarta.
- Badan Pusat Statistik. (2013). *Data Strategis BPS*. BPS, Jakarta.
- Badan Pusat Statistik. (2013). *Indeks Pembangunan Manusia 2012*. BPS, Jakarta.
- Badan Pusat Statistik. (2013). *Penghitungan dan Analisis Kemiskinan Makro Indonesia Tahun 2012*. BPS, Jakarta.
- Badan Pusat Statistik. (2013). *Indikator Kesejahteraan Rakyat 2012*. BPS, Jakarta.
- Bakin, S., Hegland, M. and Osborne, M. (1997). Can MARS be Improved with B-Splines?. *Computational Techiques and Applications Conference*, CTAC97.
- Broxson, B. J. (2006). *The Kronecker Product*. Theses and Disertations. University of North Florida.
- Budiantara, I.N. (2006). Model Spline dengan Knots Optimal. *Jurnal Ilmu Dasar* 7(2), 77-85.
- Budiantara, I. N., Diana, R., Puhadi and Darmesto, S. (2011). Relationship Pattern of Poverty and Unemployment in Indonesia with Bayesian

- Chamidah, N., Budiantara, I.N., Sunaryo, S., and Zain, I. (2012). Designing of Child Growth Chart Based on Multiresponse Local Polynomial Modelling. *Journal of Mathematics and Statistics*, 8(3), 342-347.
- Chaudhuri, S. (2003). *Assessing Vulnerability to Poverty: Concepts, Empirical Methods and Illustrative Examples*. Department of Economics Columbia University.
- Chambers, R. (1983). *Pembangunan Desa Mulai dari Belakang*. LP3ES, Jakarta.
- Christensen, R. (1991). *Linear Models for Multivariate, Time Series and Spatial Data*. Springer-Verlag, New York.
- Christensen, R. (2011). *Plane Answer to Complex Questions: The Theory of Linear Models*. Springer-Verlag, New York.
- Crainiceanu, C.M., Ruppert, D. and Wand, M.P. (2004). Bayesian Analysis for Penalized Spline Regression using WinBugs. *Statistical Software*, 14(14), 6-14.
- Craven, P. and Wahba, G. (1979). Smoothing Noisy Data with Spline Functions. Estimating The Correct Degree of Smoothing by The Method of Generalized Cross Validation. *Numer. Math.* (31) 317-403.
- Denison, D.G.T., Mallick, B.K. and Smith. A.F.M. (1998). Bayesian MARS. *Statistics and Computing*, 8: 337-346.
- Draper, N.R. and Smith, H. (1998). *Applied Regression Analysis*. Third Edition. John Wiley & Sons.
- Eubank, R.L. (1999). *Nonparametric Regression and Spline Smoothing*. Second Edition. New York. Marcel Dekker, Inc.
- Fernandes, A. A. R., Budiantara, I. N., Otok, B. W., and Suhartono (2014). Spline Estimator for Bi-responses Nonparametric Regression Model for Longitudinal Data. *International Journal of Applied Mathematical Sciences*, Vol 8. No 113-116, pp 5653-5655.
- Foster, J., Greer, J., and Thorbecke, E. (1984). A Class of Decomposable Poverty Measures. *Econometrica*, 52 : 761-766.
- Friedman, J.H. (1990). *Some Comments on Multivariate Adaptive Regression Splines*. Seattle, WA 98195.

- Friedman, J.H. (1991). *Estimating Functions of Mixed Ordinal and Categorical Variables using Adaptive Splines*. Technical Report NO. 108, Department of Statistics, Stanford University.
- Friedman, J.H. (1991). *Multivariate Adaptive Regression Splines (with discussion)*. *Annual Statistics*. 19:1-141.
- Friedman, J.H. and Silverman, B.W. (1989). Flexible Parsimony Smoothing and Additive Modeling. *Technometrics*, 31, 3-39.
- Gujarati, D.N. (2003). *Basic Econometric*. Fourth Edition. Mc. Graw Hill, New York.
- Gupta, A. K. and Nagar, D. K. (2000). *Matrix Variate Distributions*. Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics, 104.
- Hallin, B.M., Lu, Z., and Tran, L.T. (2004). Local Linear Spatial Regression. *The Annals of Statistics*, Vol. 32, No.6, 2469-2500.
- Hardle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press, Cambridge U.K.
- Hastie, T., Tibshirani, R. and Friedman, J.H. (2008). *The Element of Statistical Learning: Data Mining, Inference and Prediction*. Springer Series in Statistics, New York.
- Hidayat, N. K. (2008). *Analisis Hubungan Komponen Indeks Pembangunan Manusia dengan Kemiskinan di Propinsi Jawa Barat*. Fakultas Pertanian, Institut Pertanian Bogor.
- Howell, J.R. (2007). *Analysis Using Smoothing Splines As Implemented In LME()* In R. Thesis. Brigham Young University.
- Ijaiya, G. T., Bello, R. A. and Ajayi, M. A. (2011). Economic Growth and Poverty Reduction in Nigeria. *International Journal of Business and Social Science*. Vol. 2 No. 15.
- Jekabsons, G. (2015). *ARESLab : Adaptive Regression Splines Toolbox for Matlab/Octave ver. 1.8.2*. Institute of Applied Computer Systems Riga Technical University.
- Johnson, R.A. and Wichern, D.W. (1992). *Applied Multivariate Statistical Analysis*. Prentice Hall, Englewood Chiffs, New Jersey.
- Kayri, M. and Zirhlioğlu, G. (2009). Kernel Smoothing Function and Choosing Bandwidth for Non-Parametric Regression Methods. *Ozean Journal of Applied Sciences*, 2(1), 49-54.

- Kim, Y.J. and Gu, C. (2004). Smoothing Spline Gaussian Regression: More Scalable Computation via Efficient Approximation. *Royal Statistical Society: Series B*, 66(2), 337-356.
- Lestari, B. (2008). Penalized Weighted Least-Square Estimator for Bivariate Nonparametric Regression Model with Correlated Errors. *Proceeding of the National Seminar on Mathematics and Statistics*, (MS'08), Airlangga University, Surabaya, pp: 83-95.
- Lestari, B., Budiantara, I.N., Sunaryo, S. and Mashuri, M. (2010). Spline Estimator in Multiresponse Nonparametric Regression Model with Unequal Correlation of Errors. *Journal of Mathematics and Statistics*, 6 (3): 327-332.
- Lestari, B., Budiantara, I.N., Sunaryo, S. and Mashuri, M. (2012). Spline Smoothing for Multi-Response Nonparametric Regression Model in Case of Heteroscedasticity of Variance. *Journal of Mathematics and Statistics*, 8 (3): 377-384.
- Lewis, P.A.W and Stevens, J.G. (1990). *Nonlinear Modeling of Time Series using Multivariate Adaptive Regression Splines (MARS)*. Naval Postgraduate School. Monterey, California.
- Liang, H. and Chen, J.D.Z. (2005). Assessment of the Esophageal Pressure in Gastroesophageal Reflux Disease by Local Regression. *Annals of Biomedical Engineering*, Vol. 33, No. 6, pp. 847-853.
- Lu, C.J., Lee, T.S. and Lian, C.M. (2010). Sales Forecasting of IT Products using A Hybrid MARS and SVR Model. *IEEE International Conference on Data Mining Workshops*.
- Madan, S. (2012). *Human Development and Poverty - A Perspective Across Indian States*. I. G. P. G. Regional Centre, Mirpur, M. D. University, Rohtak, Haryana, India.
- Milborrow, S. (2015). *Derived from MDA: MARS by Trevor Hastie and Rob Tibshirani. earth: Multivariate Adaptive Regression Spline Models, 2011*. R package <http://CRAN.R-project.org/package=earth>.
- Milborrow, S. (2015). *Notes on The Earth Package*. Website: <http://www.milbo.org/doc/earth-notes.pdf>.
- Mina, C. D. (2008). *Unearthing Poverty with MARS: Application of Multivariate Adaptive Regression Splines (MARS) in Identifying Household Poverty Correlates in The Philippines*. The Faculty of School of Statistics University Of The Philippines – Diliman.

- Montgomery, D.C., Peck, E.A. and Vining, G.G. (2012). *Introduction to Linear Regression Analysis*. Fifth Edition. John Wiley & Sons.
- Nottingham Q.J. and Cook D.F. (2001). Local Linear Regression for Estimating Time Series Data. *Computational Statistics & Data Analysis* 37, 209-217.
- Nurbaeti (2013). Pengaruh Pengangguran dan Pertumbuhan Ekonomi terhadap Indeks Pembangunan Manusia Kabupaten/Kota di Propinsi Jawa Tengah Tahun 2007-2011. *Economics Development Analysis Journal*. EDAJ 2 (3).
- Oshima, H.T. (1990). Employment Generation. The Long-Term Solution to Poverty. *Asian Development Review*. Vol. 8 No 1.
- Osinubi (2005). *Macroeconometric Analysis Of Growth, Unemployment and Poverty in Nigeria, Pakistan*. Economic and Social Review. Volume XLIII, No. 2 (Winter 2005), pp. 249-269.
- Otok, B.W. (2005). Klasifikasi Perbankan dengan Pendekatan CART dan MARS. *Jurnal Widya Manajemen & Akuntansi*, UWM Surabaya.
- Otok, B.W. (2008). Asimtotik Estimator Multivariate Adaptive Regression Splines. *Jurnal Nature Indonesia* (JNI), Riau.
- Otok, B.W. (2008). *Pendekatan Bootstrap pada Model Multivariate Adaptive Regression Splines* (MARS). Disertasi FMIPA UGM Yogyakarta.
- Otok, B.W. (2009). Konsistensi dan Asimtotik Normalitas Model Multivariate Adaptive Regression Spline (MARS) Respon Biner. *Jurnal Ilmu Dasar*, Vol. 10 No. 2 : 133-140.
- Otok, B.W., Subanar dan Guritno, S. (2006). Faktor-Faktor yang Mempengaruhi Volume Perdagangan Saham Menggunakan Multivariate Adaptive Regression Splines. *Jurnal Widya Manajemen & Akuntansi*, Vol 6, No. 3, UWM, Surabaya.
- Pudjianto, B. dan Syawie, M. (2015). Kemiskinan dan Pembangunan Manusia, Pusat Penelitian dan Pengembangan Kesejahteraan Sosial, Kementerian Sosial RI. *Sosio Informa* Vol. 1, No. 03.
- Pane, R., Budiantara, I.N., Zain, I. and Otok, B.W. (2014). Parametric and Nonparametric Estimators in Fourier Series Semiparametric Regression and Their Characteristics. *Applied Mathematical Sciences*, Vol. 8, no. 102, 5053-5064.

- Pratowo, N.I. (2011). Analisis Faktor-Faktor yang Berpengaruh Terhadap Indeks Pembangunan Manusia. *Jurnal Studi Ekonomi Indonesia*. Fakultas Ekonomi Universitas Sebelas Maret.
- Ranis, G., Stewart, F. and Ramirez, A. (2000). Economic Growth and Human Development. *World Development* Vol. 28, No. 2, pp. 197-219.
- Rencher, A. C. (2000). *Methods of Multivariate Analysis*. John Wiley & Sons, Inc. New York.
- Ratnasari, G. (2014). Pengaruh Tingkat Pengangguran Terhadap Tingkat Kemiskinan di Indonesia. *Jurnal Universitas Gunadarma*.
- Ripley, B.D. (1994). Neural Networks and Related Methods for Classification. *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 56, no 3, pp. 409–456.
- Rostow, W.W. (1960). *The Stages of Economic Growth: A Non-Communist Manifesto*. Cambridge, University Press.
- Ruliana, Budiantara, I.N. Otok, B.W. and Wibowo, W. (2015). Parameter Estimation of Nonlinear Structural Model SEM using Spline Approach. *Applied Mathematical Sciences*, Vol. 9, no. 149, 7439-7451.
- Searle, S. R., Casella, G. and McCulloch, C. E. (2006). *Variance Components*. John Wiley and Sons, Inc., Hoboken, New Jersey.
- Sephton, P. (2001). *Forecasting Recessions: Can We Do Better on MARS*. The Federal Reserve Bank of St. Louis.
- Septiana (2012). *Analisis Hubungan IPM, Kapasitas Fiskal dan Korupsi terhadap Kemiskinan di Indonesia*. Fakultas Ekonomi dan Bisnis, Universitas Diponegoro.
- Smith, P.L. (1979). Splines As a Useful and Convenient Statistical Tool. *The American Statistician*, Vol. 33. No. 2, pp 57-62.
- Son, H. H. and Kakwani, N. (2004). Measuring the Impact of Price Changes on Poverty: with applications to Thailand and Korea. *Journal of Economic Inequality*, Vol. 4, No. 2, 181-207.
- Son, H. H. and Kakwani, N. (2006). *Economic Growth and Poverty Reduction: Initial Conditions Matter*. United Nations Development Programme International Poverty Centre.
- Streeten, Paul, Burki, S.J., Haq, M., Hicks, N. and Stewart, F. (1981). *First Things First, Meeting Basic Human Needs in Developing Countries*. New York: Oxford University Press.

- Su, L. and Ullah, A. (2008). Local Polynomial Estimation of Nonparametric Simultaneous Equation Models. *Journal of Econometrics* 144, 193-218.
- Sudiarsa, I.W., Budiantara, I.N., Suhartono and Purnami, S.W. (2015). Combined Estimator Fourier Series and Spline Truncated in Multivariable Nonparametric Regression. *Applied Mathematical Sciences*, Vol. 9, no. 100, 4997 – 5010.
- Sukirno, S. (2004). Makro Ekonomi. Edisi Ketiga. Jakarta: PT. Raja Grafindo Persada.
- Sumarjan, S. (1993). *Kemiskinan (Suatu Pandangan Sosiologis)*. Makalah, Jakarta.
- Tambunan (2001). *Perekonomian Indonesia, Teori dan Temuan Empiris*. Jakarta: Ghalia Indonesia.
- Tarigan, L.V. (2014). *Analisis Kausalitas antara Indeks Pembangunan Manusia dan Kemiskinan di Indonesia*. Fakultas Ekonomi, Universitas Sumatera Utara.
- Todaro (2003). *Pembangunan Ekonomi di Dunia Ketiga*. Edisi Kedelapan. Jakarta: Erlangga.
- Todaro, M.P dan Smith, S. (2006). *Pembangunan Ekonomi*. Edisi Sembilan, Jilid 1. Jakarta: Erlangga.
- Wahba, G. (1990). *Spline Models for Observational Data*. Pennsylvania. SIAM.
- Wang, Y. (1998). Spline Smoothing Models with Correlated Errors. *Journal of the American Statistical Association*, 93, 341-348.
- Wang, Y., Guo, W. and Brown, M.B. (2000). Spline Smoothing for Bivariate Data with Application to Association Between Hormones. *Statistica Sinica* 10, 377-397.
- Welsh, A. H. and Yee, T. W. (2006). Local Regression for Vector Responses. *Journal of Statistical Planning and Inference*, 136, 3007 – 3031.
- Wibowo, W., Haryatmi, S. and Budiantara, I.N. (2012). On Multiresponse Semiparametric Regression. *Journal of Mathematics and Statistics*, Vol. 8, Issue 4, 489-499.
- Wu, H. and Zhang, J. (2006). *Nonparametric Regression Methods for Longitudinal Data Analysis*. John Wiley and Sons, Inc., Hoboken, New Jersey.

- Wulandari, I.D.A.M.I. dan Budiantara, I.N. (2014). Analisis Faktor-Faktor yang Mempengaruhi Persentase Penduduk Miskin dan Pengeluaran Perkapita Makanan di Jawa Timur menggunakan Regresi Nonparametrik Birespon Spline. *Jurnal Sains dan Seni Pomits* Vol. 3(1) : 2337-3520 (2301-928X Print).
- Xiong, R. and Meullenet, J.F. (2002). *Comparison of Logistic Regression and MARS in Modelling The Effects of Water Activity, pH and Potassium Sorbate on Growth – No Growth of Saccharomyces Cerevisiac*. Food Science Department, University of Arkansas.
- Xu, Q.S., Daszykowski, M., Walczak, B., Daeyaert, F., Jonge, M.R., Heeres, J., Koymans, L.M.H., Lewi, P.J., Vinkers, H.M., Janssen, P.A. and Massart, D.L. (2004). Multivariate Adaptive Regression Splines-Studies of HIV Reverse Transcriptase Inhibitors. *Chemometrics and Intelligent Laboratory Systems* 72, 27-34.
- Xu, Q.S., Massart, D.L., Liang, Y. and Fang, K. (2003). Two-Step Multivariate Adaptive Regression Splines for Modeling A Quantitative Relationship between Gas Chromatography Retention Indices and Molecular Descriptors. *Journal of Chromatography A*, 998, 155-167.
- Yacoub, Y. (2012). Pengaruh Tingkat Pengangguran Terhadap Tingkat Kemiskinan Kabupaten/Kota di Provinsi Kalimantan Barat. *Jurnal Ilmu Ekonomi Fakultas Ekonomi Universitas Tanjung Pura*. Vol. 8 No. 3, Hal 176-185.

Lampiran 1. Syntax untuk Membangkitkan Data Simulasi pada Percobaan Fungsi Asli MARS

```
function [Y1, Y2, X_y1, X_y2] = f_asli(n, var1, var2, rho);
% =====
%   Membangkitkan data simulasi pada fungsi asli MARS dengan
%   replikasi sebanyak 100 kali
% -----
%   Input
%   n       : ukuran sampel/observasi
%   var1    : nilai varian error untuk respon 1
%   var2    : nilai varian error untuk respon 2
%   rho     : korelasi error antara respon 1 dan respon 2
% -----
%   Output
%   Y1      : data bangkitan untuk respon 1
%   Y2      : data bangkitan untuk respon 2
%   X_y1    : data bangkitan untuk prediktor pada respon 1
%   X_y2    : data bangkitan untuk prediktor pada respon 2
% =====

if nargin < 4
    error('Input kurang.');
```

end

```
% membuat cell untuk menyimpan data hasil replikasi program
Y1 = {};
Y2 = {};
X_y1 = {};
X_y2 = {};

% melakukan replikasi pembangkitan data sebanyak 100 kali
replikasi = 0;
for i = 1 : 100

    % membangkitkan prediktor
    Xy1 = unifrnd(0,1,n,2);
    Xy2 = unifrnd(0,1,n,2);
    x1 = Xy1(:,1);
    x2 = Xy1(:,2);
    x3 = Xy2(:,1);
    x4 = Xy2(:,2);

    % contoh bentuk fungsi asli MARS
    BF11 = max(0, x1 - 0.4);
    BF12 = max(0, 0.4 - x1);
    BF13 = BF11.*max(0, x2 - 0.5);
    BF14 = BF11.*max(0, 0.5 - x2);
    BF21 = max(0, x3 - 0.45);
    BF22 = max(0, 0.45 - x3);
    BF23 = BF22.*max(0, x4 - 0.7);
    BF24 = BF22.*max(0, 0.7 - x4);
    BF1 = [BF11 BF12 BF13 BF14];
    BF2 = [BF21 BF22 BF23 BF24];
    f1 = 4 + 39.*BF1(:,1) + 20.*BF1(:,2) - 35.*BF1(:,3) - 37.*BF1(:,4);
    f2 = 5 + 37.*BF2(:,1) + 18.*BF2(:,2) - 33.*BF2(:,3) - 39.*BF2(:,4);
```

Lanjutan Lampiran 1.

```
% membangkitkan error
std1 = sqrt(var1);
std2 = sqrt(var2);
cov = rho*std1*std2;
mu = [0 0];
sigma = [var1 cov; cov var2];
er = mvnrnd(mu, sigma, n);

% membangkitkan respon
y1 = f1 + er(:,1);
y2 = f2 + er(:,2);

% mengumpulkan output data bangkitan
Y1{end+1} = y1;
Y2{end+1} = y2;
X_y1{end+1} = Xy1;
X_y2{end+1} = Xy2;

replikasi = replikasi + 1
end
return
```

Lampiran 2. Syntax untuk Membangkitkan Data Simulasi pada Kombinasi Fungsi Percobaan Trigonometri-Polinomial

```
function [Y1, Y2, X_y1, X_y2] = f_tripoli(n, var1, var2, rho);
% =====
% Membangkitkan data simulasi pada kombinasi fungsi trigonometri
% dan polinomial dengan replikasi sebanyak 100 kali
% -----
% Input
% n      : ukuran sampel/observasi
% var1   : nilai varian error untuk respon 1
% var2   : nilai varian error untuk respon 2
% rho    : korelasi error antara respon 1 dan respon 2
% -----
% Output
% Y1     : data bangkitan untuk respon 1
% Y2     : data bangkitan untuk respon 2
% X_y1   : data bangkitan untuk prediktor pada respon 1
% X_y2   : data bangkitan untuk prediktor pada respon 2
% =====

if nargin < 4
    error('Input kurang.');
```

end

% membuat cell untuk menyimpan data hasil replikasi program

```
Y1 = {};
Y2 = {};
X_y1 = {};
X_y2 = {};

% melakukan replikasi pembangkitan data sebanyak 100 kali
replikasi = 0;
for i = 1 : 100

    % membangkitkan prediktor
    Xy1 = unifrnd(0,1,n,2);
    Xy2 = unifrnd(0,1,n,2);
    x1 = Xy1(:,1);
    x2 = Xy1(:,2);
    x3 = Xy2(:,1);
    x4 = Xy2(:,2);

    % contoh bentuk fungsi kombinasi trigonometri-polinomial
    f1 = 5.*sin(2.*pi.*x1) - 15.*(0.5-x2).^2 + 6.*x1.*x2;
    f2 = 4.*sin(2.*pi.*x3) + 16.*(x4-0.5).^2 - 7.*x3.*x4;

    % membangkitkan error
    std1 = sqrt(var1);
    std2 = sqrt(var2);
    cov = rho*std1*std2;
    mu = [0 0];
    sigma = [var1 cov; cov var2];
    er = mvnrnd(mu, sigma, n);
```

Lanjutan Lampiran 2

```
% membangkitkan respon
y1 = f1 + er(:,1);
y2 = f2 + er(:,2);

% mengumpulkan output data bangkitan
Y1{end+1} = y1;
Y2{end+1} = y2;
X_y1{end+1} = Xy1;
X_y2{end+1} = Xy2;

replikasi = replikasi + 1
end
return
```

Lampiran 3. Syntax untuk Membangkitkan Data Simulasi pada Kombinasi Fungsi Percobaan Eksponensial-Trigonometri

```
function [Y1, Y2, X_y1, X_y2] = f_expotri(n, var1, var2, rho);
% =====
% Membangkitkan data simulasi pada kombinasi fungsi eksponensial
% dan trigonometri dengan replikasi sebanyak 100 kali
% -----
% Input
% n      : ukuran sampel/observasi
% var1   : nilai varian error untuk respon 1
% var2   : nilai varian error untuk respon 2
% rho    : korelasi error antara respon 1 dan respon 2
% -----
% Output
% Y1     : data bangkitan untuk respon 1
% Y2     : data bangkitan untuk respon 2
% X_y1   : data bangkitan untuk prediktor pada respon 1
% X_y2   : data bangkitan untuk prediktor pada respon 2
% =====

if nargin < 4
    error('Input kurang.');
```

```
end

% membuat cell untuk menyimpan data hasil replikasi program
Y1 = {};
Y2 = {};
X_y1 = {};
X_y2 = {};

% melakukan replikasi pembangkitan data sebanyak 100 kali
replikasi = 0;
for i = 1 : 100

    % membangkitkan prediktor
    Xy1 = unifrnd(0,1,n,2);
    Xy2 = unifrnd(0,1,n,2);
    x1 = Xy1(:,1);
    x2 = Xy1(:,2);
    x3 = Xy2(:,1);
    x4 = Xy2(:,2);

    % contoh bentuk fungsi kombinasi eksponensial-trigonometri
    f1 = 6.*sin(3.*pi.*x1) + 3.*exp(-9.*x2)+ 7.*x1.*x2;
    f2 = 5.*sin(3.*pi.*x3) - 2.*exp(-8.*x4)+ 8.*x3.*x4;

    % membangkitkan error
    std1 = sqrt(var1);
    std2 = sqrt(var2);
    cov = rho*std1*std2;
    mu = [0 0];
    sigma = [var1 cov; cov var2];
    er = mvnrnd(mu, sigma, n);
```

Lanjutan Lampiran 3

```
% membangkitkan respon
y1 = f1 + er(:,1);
y2 = f2 + er(:,2);

% mengumpulkan output data bangkitan
Y1{end+1} = y1;
Y2{end+1} = y2;
X_y1{end+1} = Xy1;
X_y2{end+1} = Xy2;

replikasi = replikasi + 1
end
return
```

Lampiran 4. Syntax untuk Membangkitkan Data Simulasi pada Kombinasi Fungsi Percobaan Polinomial-Eksponensial

```
function [Y1, Y2, X_y1, X_y2] = f_poliexpo(n, var1, var2, rho);
% =====
% Membangkitkan data simulasi pada kombinasi fungsi polinomial
% dan eksponensial dengan replikasi sebanyak 100 kali
% -----
% Input
% n      : ukuran sampel/observasi
% var1   : nilai varian error untuk respon 1
% var2   : nilai varian error untuk respon 2
% rho    : korelasi error antara respon 1 dan respon 2
% -----
% Output
% Y1     : data bangkitan untuk respon 1
% Y2     : data bangkitan untuk respon 2
% X_y1   : data bangkitan untuk prediktor pada respon 1
% X_y2   : data bangkitan untuk prediktor pada respon 2
% =====

if nargin < 4
    error('Input kurang.');
```

```
end

% membuat cell untuk menyimpan data hasil replikasi program
Y1 = {};
Y2 = {};
X_y1 = {};
X_y2 = {};

% melakukan replikasi pembangkitan data sebanyak 100 kali
replikasi = 0;
for i = 1 : 100

    % membangkitkan prediktor
    Xy1 = unifrnd(0,1,n,2);
    Xy2 = unifrnd(0,1,n,2);
    x1 = Xy1(:,1);
    x2 = Xy1(:,2);
    x3 = Xy2(:,1);
    x4 = Xy2(:,2);

    % contoh bentuk fungsi kombinasi polinomial-eksponensial
    f1 = 15.*(0.5-x1).^2 + 5.*exp(-12.*x2) + 8.*x1.*x2;
    f2 = 16.*(x3-0.7).^2 + 4.*exp(-7.*x4) + 7.*x3.*x4;

    % membangkitkan error
    std1 = sqrt(var1);
    std2 = sqrt(var2);
    cov = rho*std1*std2;
    mu = [0 0];
    sigma = [var1 cov; cov var2];
    er = mvnrnd(mu, sigma, n);
```

Lanjutan Lampiran 4

```
% membangkitkan respon
y1 = f1 + er(:,1);
y2 = f2 + er(:,2);

% mengumpulkan output data bangkitan
Y1{end+1} = y1;
Y2{end+1} = y2;
X_y1{end+1} = Xy1;
X_y2{end+1} = Xy2;

replikasi = replikasi + 1
end
return
```


Lampiran 5. Syntax Pemodelan MARS Birespon Kontinu pada Data Simulasi

```
function [output] = simulasi_biresponmars(Y1, Y2, X_y1, X_y2);
% =====
% Memodelkan MARS birespon kontinu pada data simulasi
% yang dibangkitkan dengan replikasi sebanyak 100 kali
% -----
% Input
% Y1 : data bangkitan respon 1
% Y2 : data bangkitan respon 2
% X_y1 : data bangkitan prediktor untuk respon 1
% X_y2 : data bangkitan prediktor untuk respon 2
% -----
% Output
% beta_y1 : koefisien regresi MARS birespon untuk respon 1
% beta_y2 : koefisien regresi MARS birespon untuk respon 2
% knot_y1 : titik knot pada model MARS birespon untuk respon 1
% knot_y2 : titik knot pada model MARS birespon untuk respon 2
% RMSE_y1 : root mean square error (RMSE) untuk respon 1
% RMSE_y2 : root mean square error (RMSE) untuk respon 2
% RMSE_birespon : root mean square error (RMSE) model birespon
% R2_y1 : koefisien determinasi (R-square) untuk respon 1
% R2_y2 : koefisien determinasi (R-square) untuk respon 2
% R2_birespon : koefisien determinasi (R-square) model birespon
% est_var_erry1 : estimasi varian error model untuk respon 1
% est_var_erry2 : estimasi varian error model untuk respon 2
% est_korelasi : estimasi korelasi antar error pada model
% model_y1 : model MARS birespon untuk respon 1
% model_y2 : model MARS birespon untuk respon 2
% =====

% membuat kolom dan cell untuk menyimpan hasil replikasi program
beta_y1 = {};
beta_y2 = {};
knot_y1 = {};
knot_y2 = {};
RMSE_y1 = [];
RMSE_y2 = [];
RMSE_birespon = [];
R2_y1 = [];
R2_y2 = [];
R2_birespon = [];
est_var_erry1 = [];
est_var_erry2 = [];
est_korelasi = [];
model_y1 = {};
model_y2 = {};

% melakukan running program pada data bangkitan hasil replikasi
replikasi = 0;
for i = 1 : 100

% mengambil data respon dan prediktor
y1 = Y1{i};
y2 = Y2{i};
ygab = [y1;y2];
```

Lanjutan Lampiran 5

```
Xy1 = X_y1{i};
Xy2 = X_y2{i};

% banyaknya observasi
n = length(y1); % karena jumlah observasi Y1 dan Y2 sama

% input maximum basis fungsi dan maximum interaksi pada program
mfunc = [9,11,13,15,17,19,21];
maxitr = [1,2];
minobs = [1];
endspan = [];
threshold = 0.01;
lmf=length(mfunc);
lmi=length(maxitr);

output_olsy1 = {};
for i = 1:lmf
    for j = 1:lmi
        output_olsy1{end+1} = marsbuild(Xy1, y1, mfunc(i), ...
            maxitr(j),minobs, endspan, threshold);
    end
end

gcv_olsy1 = [];
hitung = 1;
for i = 1:lmf
    for j = 1:lmi
        gcv_olsy1(end+1)= output_olsy1{hitung}.model.GCV;
        hitung = hitung + 1;
    end
end

[~, ind_olsy1] = min(gcv_olsy1);
model_olsy1 = output_olsy1{ind_olsy1}.model;
basisy1 = output_olsy1{ind_olsy1}.B;
output_olsy2 = {};

for i=1:lmf
    for j=1:lmi
        output_olsy2{end+1} = marsbuild(Xy2, y2, mfunc(i), ...
            maxitr(j),minobs, endspan, threshold);
    end
end

gcv_olsy2 = [];
hitung = 1;
for i=1:lmf
    for j=1:lmi
        gcv_olsy2(end+1)= output_olsy2{hitung}.model.GCV;
        hitung = hitung + 1;
    end
end
```

Lanjutan Lampiran 5

```
[~, ind_olsy2] = min(gcv_olsy2);
model_olsy2 = output_olsy2{ind_olsy2}.model;
basisy2 = output_olsy2{ind_olsy2}.B;

% mengestimasi matriks bobot
y1hatols = arespredict(model_olsy1,Xy1);
y2hatols = arespredict(model_olsy2,Xy2);
ygabhatols = [y1hatols;y2hatols];
errols_y1 = y1 - y1hatols;
errols_y2 = y2 - y2hatols;
errols = [errols_y1 errols_y2];
sigmahat = (errols'*errols)/n;
W = kron(sigmahat,eye(n,n));

% membuat model birespon
basis = blkdiag(basisy1,basisy2);
basisw = basis'*inv(W);
coefs_w = (basisw*basis)\(basisw*ygab);
coefs_wy1 = coefs_w(1:size(basisy1,2));
coefs_wy2 = coefs_w(size(basisy1,2)+1:end);
y1hatw = basisy1*coefs_wy1;
y2hatw = basisy2*coefs_wy2;
ygabhatw = [y1hatw;y2hatw];
errwy1 = (y1 - y1hatw);
errwy2 = (y2 - y2hatw);
errw = [errwy1 errwy2];
covarians = (errw'*errw)/n;
var_errwy1 = covarians(1,1);
var_errwy2 = covarians(2,2);
cov_errw12 = covarians(1,2);
rho_w = cov_errw12/sqrt(var_errwy1*var_errwy2);
mse_wy1 = mean((y1 - y1hatw).^2);
mse_wy2 = mean((y2 - y2hatw).^2);
mse_wygab = mean((ygab - ygabhatw).^2);
rmse_wy1 = sqrt(mse_wy1);
rmse_wy2 = sqrt(mse_wy2);
rmse_wygab = sqrt(mse_wygab);
R2_wy1 = 1 - mse_wy1/var(y1,1);
R2_wy2 = 1 - mse_wy2/var(y2,1);
y1_bar = mean(y1).*ones(n,1);
y2_bar = mean(y2).*ones(n,1);
ygab_bar = [y1_bar;y2_bar];
R2_wygab = 1 - (sum((ygab - ygabhatw).^2))/(sum((ygab - ygab_bar).^2));
model_wy1 = model_olsy1;
model_wy1.coefs = coefs_wy1;
model_wy1.MSE = mse_wy1;
model_wy2 = model_olsy2;
model_wy2.coefs = coefs_wy2;
model_wy2.MSE = mse_wy2;

% mengumpulkan output program
beta_y1{end+1} = coefs_wy1;
beta_y2{end+1} = coefs_wy2;
```

Lanjutan Lampiran 5

```
knot_y1{end+1} = model_wy1.knotsites;
knot_y2{end+1} = model_wy2.knotsites;
RMSE_y1(end+1) = rmse_wy1;
RMSE_y2(end+1) = rmse_wy2;
RMSE_birespon(end+1) = rmse_wygab;
R2_y1(end+1) = R2_wy1;
R2_y2(end+1) = R2_wy2;
R2_birespon(end+1) = R2_wygab;
est_var_erry1(end+1) = var_errwy1;
est_var_erry2(end+1) = var_errwy2;
est_korelasi(end+1) = rho_w;
model_y1{end+1} = model_wy1;
model_y2{end+1} = model_wy2;
replikasi = replikasi + 1
end

%menyajikan output program
output.beta_y1 = beta_y1;
output.beta_y2 = beta_y2;
output.knot_y1 = knot_y1;
output.knot_y2 = knot_y2;
output.RMSE_y1 = RMSE_y1;
output.RMSE_y2 = RMSE_y2;
output.RMSE_birespon = RMSE_birespon;
output.R2_y1 = R2_y1;
output.R2_y2 = R2_y2;
output.R2_birespon = R2_birespon;
output.est_var_erry1 = est_var_erry1;
output.est_var_erry2 = est_var_erry2;
output.est_korelasi = est_korelasi;
output.model_y1 = model_y1;
output.model_y2 = model_y2;
return
```

Lampiran 6. Syntax untuk Menghitung Rasio *RMSEP* pada Data Simulasi

```
function [output] = rasio_RMSEP_simulasi(Y1,Y2,X_y1,X_y2,k,rho);
% =====
% Menghitung rasio Root Mean Square Error Prediction (RMSEP)model
% MARS birespon kontinu pada data simulasi yang dibangkitkan
% dengan replikasi sebanyak 100 kali
% -----
% Input
% Y1 : data bangkitan respon 1
% Y2 : data bangkitan respon 2
% X_y1 : data bangkitan prediktor untuk respon 1
% X_y2 : data bangkitan prediktor untuk respon 2
% k : banyaknya fold
% rho : korelasi error antar respon 1 dan respon 2
% -----
% Output
% RMSEP_olsy1 : Nilai RMSEP untuk respon 1 (tanpa bobot/korelasi)
% RMSEP_olsy2 : Nilai RMSEP untuk respon 2 (tanpa bobot/korelasi)
% RMSEP_olsygab : Nilai RMSEP untuk gabungan respon(tanpa bobot)
% RMSEP_y1 : Nilai RMSEP birespon untuk respon 1
% RMSEP_y2 : Nilai RMSEP birespon untuk respon 2
% RMSEP_birespon : Nilai RMSEP birespon
% rasio_RMSEP_y1 : Nilai rasio RMSEP birespon untuk respon 1
% rasio_RMSEP_y2 : Nilai rasio RMSEP birespon untuk respon 1
% rasio_RMSEP_birespon : Nilai rasio RMSEP birespon
% keterangan:
% rasio_RMSEP diperoleh dari pembagian nilai RMSEP (tanpa bobot)
% dengan nilai RMSEP birespon (dengan melibatkan bobot)
% =====

% membuat cell untuk menyimpan hasil replikasi program
RMSEP_ols_y1 = [];
RMSEP_ols_y2 = [];
RMSEP_ols_ygab = [];
RMSEP_y1 = [];
RMSEP_y2 = [];
RMSEP_birespon = [];
rasio_RMSEP_y1 = [];
rasio_RMSEP_y2 = [];
rasio_RMSEP_birespon = [];

% melakukan running program pada data bangkitan hasil replikasi
replikasi = 0;
for i = 1 : 100

% memanggil data respon dan prediktor
y1 = Y1{i};
y2 = Y2{i};
ygab = [y1;y2];
Xy1 = X_y1{i};
Xy2 = X_y2{i};

% banyaknya observasi
n = length(y1); % karena jumlah observasi Y1 dan Y2 sama
```

Lanjutan Lampiran 6

```
% menentukan k-fold
if (nargin < 5) || (isempty(k))
    k = 10;
end
if k < 2
    error('k tidak boleh lebih kecil dari 2.');
```

```
end
if k > n
    error('k tidak boleh lebih besar dari jumlah observasi (n).');
```

```
end

% membagi data kedalam k-fold (k-grup)
minsize = floor(n / k);
remainder = n - minsize * k;
sizes = zeros(k, 1);
for i = 1 : k
    sizes(i) = minsize;
    if remainder > 0
        sizes(i) = sizes(i) + 1;
        remainder = remainder - 1;
    end
end
offsets = ones(k, 1);
for i = 2 : k
    offsets(i) = offsets(i-1) + sizes(i-1);
end

% Membuat kolom nilai RMSEP data testing sebanyak k-fold
RMSEP_olsy1 = NaN(k,1);
RMSEP_olsy2 = NaN(k,1);
RMSEP_olsygab = NaN(k,1);
RMSEP_wy1 = NaN(k,1);
RMSEP_wy2 = NaN(k,1);
RMSEP_wygab = NaN(k,1);

% membuat data training dan data testing berdasarkan k-fold
ind = 1 : n;
for i = 1 : k
    disp(['Fold #' num2str(i)]);

    % membuat data training
    Xtry1 = zeros(n-sizes(k-i+1), size(Xy1,2));
    Xtry2 = zeros(n-sizes(k-i+1), size(Xy2,2));
    Ytr1 = zeros(n-sizes(k-i+1), 1);
    Ytr2 = zeros(n-sizes(k-i+1), 1);
    currsize = 0;
    for j = 1 : k
        if k-i+1 ~= j
            Xtry1(currsize+1 : currsize+1+sizes(j)-1, :) = ...
                Xy1(ind(offsets(j)): offsets(j)+sizes(j)-1, :));
            Xtry2(currsize+1 : currsize+1+sizes(j)-1, :) = ...
                Xy2(ind(offsets(j)): offsets(j)+sizes(j)-1, :));
            Ytr1(currsize+1 : currsize+1+sizes(j)-1, :) = ...
                y1(ind(offsets(j)): offsets(j)+sizes(j)-1, :));
            Ytr2(currsize+1 : currsize+1+sizes(j)-1, :) = ...
                y2(ind(offsets(j)): offsets(j)+sizes(j)-1, :);
```

Lanjutan Lampiran 6

```
        currsize = currsize + sizes(j);
    end
end
Ytrgab = [Ytr1;Ytr2];

% membuat data testing
Xtsty1=Xy1(ind(offsets(k-i+1):offsets(k-i+1)+sizes(k-i+1)-1),:);
Xtsty2=Xy2(ind(offsets(k-i+1):offsets(k-i+1)+sizes(k-i+1)-1),:);
Ytst1 = y1(ind(offsets(k-i+1):offsets(k-i+1)+sizes(k-i+1)-1),:);
Ytst2 = y2(ind(offsets(k-i+1):offsets(k-i+1)+sizes(k-i+1)-1),:);
Ytstgab = [Ytst1;Ytst2];

% membuat model tanpa bobot
mfunc = 21;
maxitr = 2;
minobs = [];
endspan = [];
threshold = 0.0001;
output_olsy1 = marsbuild(Xtry1, Ytr1, mfunc, maxitr, minobs,...
    endspan, threshold);
output_olsy2 = marsbuild(Xtry2, Ytr2, mfunc, maxitr, minobs,...
    endspan, threshold);
model_olsy1 = output_olsy1.model;
model_olsy2 = output_olsy2.model;

% menghitung RMSEP model tanpa bobot
y1hatols_tst = arespredict(model_olsy1,Xtsty1);
y2hatols_tst = arespredict(model_olsy2,Xtsty2);
yhatgabols_tst = [y1hatols_tst;y2hatols_tst];
olsmsepy1 = mean((Ytst1 - y1hatols_tst).^2);
olsmsepy2 = mean((Ytst2 - y2hatols_tst).^2);
olsmsepygab = mean((Ytstgab - yhatgabols_tst).^2);
olsrmsepy1 = sqrt(olsmsepy1);
olsrmsepy2 = sqrt(olsmsepy2);
olsrmsepygab = sqrt(olsmsepygab);
RMSEP_olsy1(i) = olsrmsepy1;
RMSEP_olsy2(i) = olsrmsepy2;
RMSEP_olsygab(i) = olsrmsepygab;

% mengestimasi matriks bobot
y1hatols_tr = arespredict(model_olsy1,Xtry1);
y2hatols_tr = arespredict(model_olsy2,Xtry2);
ygabhatols_tr = [y1hatols_tr;y2hatols_tr];
errols_y1 = Ytr1 - y1hatols_tr;
errols_y2 = Ytr2 - y2hatols_tr;
errols = [errols_y1 errols_y2];
ntr = length(errols);
sigmahats = (errols'*errols)/ntr;
if rho == 0
    sigmahats = diag(diag(sigmahats));
end
W = kron(sigmahats,eye(ntr,ntr));
```

Lanjutan Lampiran 6

```
% membuat model birespon
basisy1 = output_olsy1.B;
basisy2 = output_olsy2.B;
basisygab = blkdiag(basisy1,basisy2);
basisw = basisygab'*inv(W);
coefs_w = (basisw*basisygab)\(basisw*Ytrgab);
coefs_wy1 = coefs_w(1:size(basisy1,2));
coefs_wy2 = coefs_w(size(basisy1,2)+1:end);
model_wy1 = model_olsy1;
model_wy1.coefs = coefs_wy1;
model_wy1.MSE = sigmahats(1,1);
model_wy2 = model_olsy2;
model_wy2.coefs = coefs_wy2;
model_wy2.MSE = sigmahats(2,2);

% menghitung RMSEP model birespon
y1hatwls_tst = arespredict(model_wy1,Xtsty1);
y2hatwls_tst = arespredict(model_wy2,Xtsty2);
yhatgabwls_tst = [y1hatwls_tst;y2hatwls_tst];
wlsmsepy1 = mean((Ytst1 - y1hatwls_tst).^2);
wlsmsepy2 = mean((Ytst2 - y2hatwls_tst).^2);
wlsmsepygab = mean((Ytstgab - yhatgabwls_tst).^2);
wlsrmsepy1 = sqrt(wlsmsepy1);
wlsrmsepy2 = sqrt(wlsmsepy2);
wlsrmsepygab = sqrt(wlsmsepygab);
RMSEP_wy1(i) = wlsrmsepy1;
RMSEP_wy2(i) = wlsrmsepy2;
RMSEP_wygab(i) = wlsrmsepygab;
end

%menghitung rata-rata RMSEP model tanpa bobot dari k fold
avgRMSEP_olsy1 = mean(RMSEP_olsy1);
avgRMSEP_olsy2 = mean(RMSEP_olsy2);
avgRMSEP_olsygab = mean(RMSEP_olsygab);

% menghitung rata-rata RMSEP model birespon dari k fold
avgRMSEP_wy1 = mean(RMSEP_wy1);
avgRMSEP_wy2 = mean(RMSEP_wy2);
avgRMSEP_wygab = mean(RMSEP_wygab);

% menghitung rasio rata-rata RMSEP (tanpa bobot/birespon)
rasio_RMSEPy1 = avgRMSEP_olsy1/avgRMSEP_wy1;
rasio_RMSEPy2 = avgRMSEP_olsy2/avgRMSEP_wy2;
rasio_RMSEPygab = avgRMSEP_olsygab/avgRMSEP_wygab;

% mengumpulkan output program
RMSEP_ols_y1(end+1) = avgRMSEP_olsy1;
RMSEP_ols_y2(end+1) = avgRMSEP_olsy2;
RMSEP_ols_ygab(end+1) = avgRMSEP_olsygab;
RMSEP_y1(end+1) = avgRMSEP_wy1;
RMSEP_y2(end+1) = avgRMSEP_wy2;
RMSEP_birespon(end+1) = avgRMSEP_wygab;
rasio_RMSEP_y1(end+1) = rasio_RMSEPy1;
rasio_RMSEP_y2(end+1) = rasio_RMSEPy2;
rasio_RMSEP_birespon(end+1)= rasio_RMSEPygab;
```


Lanjutan Lampiran 6

```
% replikasi
replikasi = replikasi + 1
end

% menyajikan output progrm
output.RMSEP_olsy1 = RMSEP_ols_y1;
output.RMSEP_olsy2 = RMSEP_ols_y2;
output.RMSEP_olsygab = RMSEP_ols_ygab;
output.RMSEP_y1 = RMSEP_y1;
output.RMSEP_y2 = RMSEP_y2;
output.RMSEP_birespon = RMSEP_birespon;
output.rasio_RMSEP_y1 = rasio_RMSEP_y1;
output.rasio_RMSEP_y2 = rasio_RMSEP_y2;
output.rasio_RMSEP_birespon = rasio_RMSEP_birespon;
return
```

Lampiran 7. Syntax Pemodelan Persentase Penduduk Miskin dan IPM Menggunakan Estimator MARS Birespon Kontinu

```
function [output] = modelmarsbirespon_kemiskinan_ipm(Y, X);
% =====
% Memodelkan MARS birespon kontinu pada data riil:
% "Pemodelan Persentase Penduduk Miskin dan IPM dengan Estimator
% MARS Birespon Kontinu"
% -----
% Input
% Y      : birespon (y1 = persentase penduduk miskin, y2 = IPM)
% X      : prediktor (x1 = pertumbuhan ekonomi, x2 = angka beban
%               ketergantungan, x3 = persentase pengeluaran perkapita
%               untuk makanan, x4 = tingkat pengangguran terbuka)
% -----
% Output
% beta_y1 : koefisien regresi MARS birespon untuk respon miskin
% beta_y2 : koefisien regresi MARS birespon untuk respon IPM
% knot_y1 : titik knot model MARS birespon untuk respon miskin
% knot_y2 : titik knot pada model MARS birespon untuk respon IPM
% rmse_y1 : root mean square error (RMSE) untuk respon miskin
% rmse_y2 : root mean square error (RMSE) untuk respon IPM
% rmse_birespon : root mean square error (RMSE) model birespon
% R2_y1      : koefisien determinasi (R-square) untuk respon miskin
% R2_y2      : koefisien determinasi (R-square) untuk respon IPM
% R2_birespon : koefisien determinasi (R-square) model birespon
% model_y1   : model MARS birespon untuk respon miskin
% model_y2   : model MARS birespon untuk respon IPM
% =====

y1 = Y(:,1);
y2 = Y(:,2);
ygab = [y1;y2];
n = length(X); % banyaknya observasi

% input maximum basis fungsi dan maximum interaksi pada program
mfunc = [9,11,13,15,17,19,21];
maxitr = [1,2,3];
minobs = [1];
endspan = [];
threshold = 0.0001;
lmf=length(mfunc);
lmi=length(maxitr);

output_olsy1 = {};
for i = 1:lmf
    for j = 1:lmi
        output_olsy1{end+1} = marsbuild(X, y1, mfunc(i), ...
            maxitr(j),minobs, endspan, threshold);
    end
end

gcv_olsy1 = [];
hitung = 1;
for i = 1:lmf
    for j = 1:lmi
        gcv_olsy1(end+1)= output_olsy1{hitung}.model.GCV;
```

Lanjutan Lampiran 7

```
        hitung = hitung + 1;
    end
end
[~, ind_olsy1] = min(gcv_olsy1);
model_olsy1 = output_olsy1{ind_olsy1}.model;
basisy1 = output_olsy1{ind_olsy1}.B;

output_olsy2 = {};
for i=1:lmf
    for j=1:lmi
        output_olsy2{end+1} = marsbuild(X, y2, mfunc(i), ...
            maxitr(j), minobs, endspar, threshold);
    end
end

gcv_olsy2 = [];
hitung = 1;
for i=1:lmf
    for j=1:lmi
        gcv_olsy2(end+1) = output_olsy2{hitung}.model.GCV;
        hitung = hitung + 1;
    end
end
[~, ind_olsy2] = min(gcv_olsy2);
model_olsy2 = output_olsy2{ind_olsy2}.model;
basisy2 = output_olsy2{ind_olsy2}.B;

% mengestimasi matriks bobot
y1hatols = arespredict(model_olsy1,X);
y2hatols = arespredict(model_olsy2,X);
ygabhatols = [y1hatols;y2hatols];
errols_y1 = y1 - y1hatols;
errols_y2 = y2 - y2hatols;
errols = [errols_y1 errols_y2];
sigmahat = (errols'*errols)/n;
W = kron(sigmahat,eye(n,n));

% membuat model birespon
basis = blkdiag(basisy1,basisy2);
basisw = basis'*inv(W);
coefs_w = (basisw*basis)\(basisw*ygab);
coefs_wy1 = coefs_w(1:size(basisy1,2));
coefs_wy2 = coefs_w(size(basisy1,2)+1:end);
y1hatw = basisy1*coefs_wy1;
y2hatw = basisy2*coefs_wy2;
ygabhatw = [y1hatw;y2hatw];
mse_wy1 = mean((y1 - y1hatw).^2);
mse_wy2 = mean((y2 - y2hatw).^2);
mse_wygab = mean((ygab - ygabhatw).^2);
rmse_wy1 = sqrt(mse_wy1);
rmse_wy2 = sqrt(mse_wy2);
rmse_wygab = sqrt(mse_wygab);
R2_wy1 = 1 - mse_wy1/var(y1,1);
R2_wy2 = 1 - mse_wy2/var(y2,1);
```

Lanjutan Lampiran 7

```
y1_bar = mean(y1).*ones(n,1);
y2_bar = mean(y2).*ones(n,1);
ygab_bar = [y1_bar;y2_bar];
R2_wygab = 1-(sum((ygab-ygabhatw).^2))/(sum((ygab-ygab_bar).^2));
model_wy1 = model_olsy1;
model_wy1.coefs = coefs_wy1;
model_wy1.MSE = mse_wy1;
model_wy2 = model_olsy2;
model_wy2.coefs = coefs_wy2;
model_wy2.MSE = mse_wy2;

%menyajikan output program
output.beta_y1 = coefs_wy1;
output.beta_y2 = coefs_wy2;
output.knot_y1 = model_wy1.knotsites;
output.knot_y2 = model_wy2.knotsites;
output.rmse_y1 = rmse_wy1;
output.rmse_y2 = rmse_wy2;
output.rmse_birespon = rmse_wygab;
output.R2_y1 = R2_wy1;
output.R2_y2 = R2_wy2;
output.R2_birespon = R2_wygab;
output.model_y1 = model_wy1;
output.model_y2 = model_wy2;
return
```

Lampiran 8. Syntax untuk Menghitung Nilai *RMSECV* pada Pemodelan Persentase Penduduk Miskin dan IPM Menggunakan Estimator MARS Birespon Kontinu

```
function [output] = RMSECV_Pemodelan_Kemiskinan_IPM(Y, X);
% =====
% Menghitung Root Mean Square Error of Cross Validation (RMSECV)
% pada "Pemodelan Persentase Penduduk Miskin dan IPM dengan
% estimator MARS Birespon Kontinu"
% -----
% Input
% Y      : birespon (y1 = persentase penduduk miskin, y2 = IPM)
% X      : prediktor (x1 = pertumbuhan ekonomi, x2 = angka beban
%               ketergantungan, x3 = persentase pengeluaran perkapita
%               untuk makanan, x4 = tingkat pengangguran terbuka)
% -----
% Output
% RMSEPCV_olsy1 : Nilai RMSEPCV untuk kemiskinan (tanpa korelasi)
% RMSEPCV_olsy2 : Nilai RMSEPCV untuk IPM (tanpa korelasi)
% RMSEPCV_olsygab : Nilai RMSEPCV (miskin & IPM) (tanpa korelasi)
% RMSEPCV_y1 : Nilai RMSEPCV birespon untuk kemiskinan
% RMSEPCV_y2 : Nilai RMSEPCV birespon untuk IPM
% RMSEPCV_birespon : Nilai RMSEPCV birespon (miskin & IPM)
% =====

y1 = Y(:,1);
y2 = Y(:,2);
ygab = [y1;y2];
n = length(X); % banyaknya observasi

% membagi data kedalam k-fold (k-grup)
k = n;
minsize = floor(n / k);
remainder = n - minsize * k;
sizes = zeros(k, 1);
for i = 1 : k
    sizes(i) = minsize;
    if remainder > 0
        sizes(i) = sizes(i) + 1;
        remainder = remainder - 1;
    end
end
offsets = ones(k, 1);
for i = 2 : k
    offsets(i) = offsets(i-1) + sizes(i-1);
end

% Membuat kolom nilai RMSEP data testing sebanyak k-fold
RMSEP_olsy1 = NaN(k,1);
RMSEP_olsy2 = NaN(k,1);
RMSEP_olsygab = NaN(k,1);
RMSEP_wy1 = NaN(k,1);
RMSEP_wy2 = NaN(k,1);
RMSEP_wygab = NaN(k,1);
```

Lanjutan Lampiran 8

```
% membuat data training dan data testing berdasarkan k-fold
ind = 1 : n;
for i = 1 : k
    disp(['Fold #' num2str(i)]);

    % membuat data training
    Xtr = zeros(n-sizes(k-i+1), size(X,2));
    Ytr1 = zeros(n-sizes(k-i+1), 1);
    Ytr2 = zeros(n-sizes(k-i+1), 1);
    currsz = 0;
    for j = 1 : k
        if k-i+1 ~= j
            Xtr(currsz+1 : currsz+1+sizes(j)-1, :) = ...
                X(ind(offsets(j): offsets(j)+sizes(j)-1), :);
            Ytr1(currsz+1 : currsz+1+sizes(j)-1, :) = ...
                y1(ind(offsets(j): offsets(j)+sizes(j)-1), :);
            Ytr2(currsz+1 : currsz+1+sizes(j)-1, :) = ...
                y2(ind(offsets(j): offsets(j)+sizes(j)-1), :);
            currsz = currsz + sizes(j);
        end
    end
    Ytrgab = [Ytr1;Ytr2];

    % membuat data testing
    Xtst = X(ind(offsets(k-i+1):offsets(k-i+1)+sizes(k-i+1)-1), :);
    Ytst1 = y1(ind(offsets(k-i+1):offsets(k-i+1)+sizes(k-i+1)-1), :);
    Ytst2 = y2(ind(offsets(k-i+1):offsets(k-i+1)+sizes(k-i+1)-1), :);
    Ytstgab = [Ytst1;Ytst2];
    % membuat model tanpa bobot
    mfunc = 21;
    maxitr = 3;
    minobs = [];
    endspan = [];
    threshold = 0.0001;
    output_olsy1 = marsbuild(Xtr, Ytr1, mfunc, maxitr, minobs, ...
        endspan, threshold);
    output_olsy2 = marsbuild(Xtr, Ytr2, mfunc, maxitr, minobs, ...
        endspan, threshold);
    model_olsy1 = output_olsy1.model;
    model_olsy2 = output_olsy2.model;

    % menghitung RMSEP model tanpa bobot
    y1hatols_tst = arespredict(model_olsy1,Xtst);
    y2hatols_tst = arespredict(model_olsy2,Xtst);
    yhatgabols_tst = [y1hatols_tst;y2hatols_tst];
    olsmsepy1 = mean((Ytst1 - y1hatols_tst).^2);
    olsmsepy2 = mean((Ytst2 - y2hatols_tst).^2);
    olsmsepygab = mean((Ytstgab - yhatgabols_tst).^2);
    olsmsepy1 = sqrt(olsmsepy1);
    olsmsepy2 = sqrt(olsmsepy2);
    olsmsepygab = sqrt(olsmsepygab);
    RMSEP_olsy1(i) = olsmsepy1;
    RMSEP_olsy2(i) = olsmsepy2;
    RMSEP_olsygab(i) = olsmsepygab;
```

Lanjutan Lampiran 8

```
% mengestimasi matriks bobot
y1hatols_tr = arespredict(model_olsy1,Xtr);
y2hatols_tr = arespredict(model_olsy2,Xtr);
ygabhatols_tr = [y1hatols_tr;y2hatols_tr];
errols_y1 = Ytr1 - y1hatols_tr;
errols_y2 = Ytr2 - y2hatols_tr;
errols = [errols_y1 errols_y2];
ntr = length(errols);
sigmahats = (errols'*errols)/ntr;
W = kron(sigmahats,eye(ntr,ntr));

% membuat model birespon
basisy1 = output_olsy1.B;
basisy2 = output_olsy2.B;
basisygab = blkdiag(basisy1,basisy2);
basisw = basisygab'*inv(W);
coefs_w = (basisw*basisygab)\(basisw*Ytrgab);
coefs_wy1 = coefs_w(1:size(basisy1,2));
coefs_wy2 = coefs_w(size(basisy1,2)+1:end);
model_wy1 = model_olsy1;
model_wy1.coefs = coefs_wy1;
model_wy1.MSE = sigmahats(1,1);
model_wy2 = model_olsy2;
model_wy2.coefs = coefs_wy2;
model_wy2.MSE = sigmahats(2,2);

% menghitung RMSEP model birespon
y1hatwls_tst = arespredict(model_wy1,Xtst);
y2hatwls_tst = arespredict(model_wy2,Xtst);
yhatgabwls_tst = [y1hatwls_tst;y2hatwls_tst];
wlsmsepy1 = mean((Ytst1 - y1hatwls_tst).^2);
wlsmsepy2 = mean((Ytst2 - y2hatwls_tst).^2);
wlsmsepygab = mean((Ytstgab - yhatgabwls_tst).^2);
wlsrmsepy1 = sqrt(wlsmsepy1);
wlsrmsepy2 = sqrt(wlsmsepy2);
wlsrmsepygab = sqrt(wlsmsepygab);
RMSEP_wy1(i) = wlsrmsepy1;
RMSEP_wy2(i) = wlsrmsepy2;
RMSEP_wygab(i) = wlsrmsepygab;
end

%menghitung output nilai RMSEPCV model tanpa bobot dari k fold
output.RMSEPCV_olsy1 = mean(RMSEP_olsy1);
output.RMSEPCV_olsy2 = mean(RMSEP_olsy2);
output.RMSEPCV_olsygab = mean(RMSEP_olsygab);

% menghitung output nilai RMSEPCV model birespon dari k fold
output.RMSEPCV_y1 = mean(RMSEP_wy1);
output.RMSEPCV_y2 = mean(RMSEP_wy2);
output.RMSEPCV_birespon = mean(RMSEP_wygab);
return
```

Lampiran 9. Syntax untuk Uji Linieritas Data

```
function [RESET df1 df2 pval] = resettest(Y,X,power,model)
% =====
% Menghitung Nilai Ramsey RESET Test
% -----
% Input
% Y      : respon
% X      : prediktor
% power  : derajat polinomial (2 atau 3) (default power = 2)
% model = 1 (untuk power = 3) (default)
%          Ybaru = a0 + a1*X + a2*Ylama^3
% model = 2 (untuk power==3) (Gujarati,2003)
%          Ybaru = a0 + a1*X + a2*Ylama^2 + a3*Ylama^3
% -----
% Output
% RESET  : Nilai Reset Test
% df1    : derajat bebas ke-1
% df2    : derajat bebas ke-2
% pval   : p-value
% =====

%default program
if nargin<4
    model = 1;
else
    model=model;
end

if nargin<3
    power = 2;
else
    power=power;
end

% Tahapan RAMSEY reset test

% step 1 (OLS)
% Y = a0+a1X+u = Z*beta1
Z = [ones(size(X,1),1) X];
beta1 = inv(Z'*Z)*Z'*Y;
Ypred_old= Z*beta1;

% mencari R_square
residual1 = abs(Y-Ypred_old);
SSResidual1 = sum(residual1.^2);
SSTotal1 = sum((Y-(ones(size(Y,1),1).*mean(Y))).^2);
Rsq_old = 1-(SSResidual1/SSTotal1);

% step 2
if model==2
if power==3
    W=[Ypred_old.^2 Ypred_old.^3];
else
    W=[Ypred_old.^2];
end
```


Lanjutan Lampiran 9

```
else
    W=[Ypred_old.^power];
End

Z2 = [Z W];
beta2 = inv(Z2'*Z2)*Z2'*Y;
Ypred_new = Z2*beta2;

% mencari R_square
residual2 = abs(Y-Ypred_new);
SSResidual2 = sum(residual2.^2);
SSTotal2 = sum((Y-(ones(size(Y,1),1).*mean(Y))).^2);
Rsq_new = 1-(SSResidual2/SSTotal2);

% step 3
%df1 = jumlah variabel bebas baru
df1 = size(W,2);
%df2=jumlah data - jumlah parameter model baru
df2 = size(X,1) - size(Z2,2);

% (Gujarati,2003)
% RESET= ((Rsq_new-Rsq_old)/df1)/((1-Rsq_new)/df2);
RESET = (df2/df1)*((SSResidual1 - SSResidual2)/SSResidual2);

% p-value
pval = 1 - fcdf(RESET, df1, df2);

end
return
```

Lampiran 10. Syntax untuk Uji Normal Multivariat Data

```
function t = qqplot(x);
% =====
% Menguji Distribusi Normal Multivariat data dengan Q-Q Plot
% -----
% Input
% x      : variabel yang akan diuji
% -----
% Output
% t       : Nilai pemeriksaan distribusi normal multivariat data
% Grafik : Grafik Q-Q Plot
% =====

[n,d] = size(x);
rata2_x = mean(x);
mean_x = ones(n,d);

for i = 1 : d
    mean_x(:,i) = rata2_x(i)*ones(n,1);
end

selisih_rata2 = (x - mean_x)';
covi_x = cov(x);

for i = 1 : n
    di(i) = selisih_rata2(:,i))'*inv(covi_x)*(selisih_rata2(:,i));
end
di = sort(di);

for i = 1 : n
    ki(i) = (i - 0.5)/n;
end

for i = 1 : n
    qi(i) = icdf('chi2',ki(i),d);
end

chisq = icdf('chi2',0.5,d);

for i = 1 : n
    if di(i) <= chisq
        ti(i) = 1;
    else
        ti(i) = 0;
    end
end
jml_ti = sum(ti);
t = jml_ti/n;

if t > 0.5
    fprintf('t = %3f\n',t);
    fprintf('karena t>0,5, maka Data Berdistribusi Normal
            Multivariat\n');
else
    fprintf('t = %3f\n',t);
```

Lanjutan Lampiran 10

```
fprintf('karena t < 0,5, maka Data Tidak Berdistribusi Normal
        Multivariat\n');
end

% membuat scatter qq-plot
scatter(di, qi, 'filled');
xlabel('di^2');
ylabel ('qi');
title('Plot Uji Q-Q Plot pada Residual Model');
set(gcf, 'Color', [1,1,1])
grid

return
```

Lampiran 11. Data Simulasi pada Fungsi Asli MARS $(n_1 = n_2 = 125; \sigma_{11} = \sigma_{22} = 1; \rho = 0,5)$

No.	y_1	y_2	x_1	x_2	x_3	x_4
1	17.2935	19.3260	0.8147	0.7021	0.8195	0.5216
2	9.6181	10.6515	0.1579	0.8329	0.0273	0.8161
3	17.6797	5.3379	0.8811	0.8107	0.4684	0.8017
4	7.5925	22.1292	0.1309	0.4003	0.9497	0.8158
5	10.2232	7.2321	0.6123	0.8616	0.4901	0.9841
6	3.9107	7.2435	0.4216	0.5183	0.1434	0.3231
7	6.3412	4.5486	0.5151	0.2485	0.4029	0.6211
8	8.3312	11.8568	0.1972	0.9473	0.6521	0.6127
9	19.7800	9.2809	0.8045	0.5982	0.5481	0.8621
10	3.8315	10.1789	0.4075	0.2807	0.6254	0.0528
11	14.5830	2.8230	0.8486	0.8772	0.2468	0.0358
12	7.9294	3.2374	0.6521	0.0394	0.1355	0.1121
13	6.9725	11.0077	0.2935	0.3716	0.6184	0.8679
14	27.3048	13.8087	0.9812	0.5160	0.6905	0.1534
15	9.5170	8.2955	0.5080	0.4561	0.1653	0.8493
16	12.7150	6.0429	0.6767	0.4297	0.3054	0.5345
17	5.9803	23.4702	0.4723	0.1082	0.9430	0.8981
18	7.5493	8.4179	0.2454	0.8973	0.3076	0.7099
19	4.8611	12.6843	0.3139	0.9566	0.6631	0.5094
20	5.3360	23.6040	0.3791	0.3560	0.9309	0.7714
21	12.5814	20.8393	0.6529	0.6653	0.9183	0.6669
22	13.2333	17.1485	0.7231	0.1691	0.7379	0.3635
23	5.5068	4.6769	0.3065	0.4200	0.4470	0.2310
24	9.3654	3.7487	0.0668	0.8779	0.1668	0.0942
25	4.0814	21.0027	0.3313	0.4535	0.8727	0.8956
26	18.8992	11.1598	0.9500	0.1649	0.0816	0.6662
27	19.9165	5.6675	0.9175	0.6959	0.4372	0.7015
28	14.9616	5.8969	0.8667	0.1427	0.1072	0.3506
29	16.2321	9.0934	0.6835	0.5828	0.1051	0.4244
30	5.9486	6.1933	0.3691	0.3233	0.4735	0.2247
31	11.2055	6.7545	0.0537	0.2233	0.1090	0.4392
32	8.5396	5.2437	0.5835	0.1507	0.4089	0.4302
33	16.3971	19.3198	0.8086	0.7704	0.8483	0.4824
34	7.6474	8.3551	0.2472	0.9760	0.5624	0.2473
35	8.4928	9.6172	0.5307	0.4798	0.5834	0.8162
36	19.4225	24.9303	0.8820	0.3711	0.9849	0.9011
37	11.9592	7.4218	0.0661	0.5845	0.4740	0.4196

Lanjutan Lampiran 11

No.	y_1	y_2	x_1	x_2	x_3	x_4
38	10.0786	19.4946	0.0927	0.3205	0.8579	0.4097
39	21.6140	7.0351	0.8766	0.5674	0.4902	0.6146
40	10.7458	7.0730	0.0044	0.1012	0.1777	0.4633
41	10.1752	7.2475	0.5787	0.2852	0.5699	0.8746
42	15.0032	11.2756	0.8451	0.0478	0.1292	0.6967
43	7.5660	18.1191	0.2342	0.2065	0.8406	0.7307
44	12.0252	9.3684	0.6544	0.8116	0.1178	0.5922
45	9.8175	7.6786	0.1552	0.5969	0.1455	0.5263
46	7.7634	1.2506	0.1503	0.9900	0.0376	0.0143
47	17.0058	10.9153	0.9071	0.8769	0.1176	0.7244
48	6.5254	8.7666	0.5034	0.8435	0.1906	0.9110
49	8.0236	17.4102	0.2064	0.2619	0.7907	0.0432
50	22.3103	6.7460	0.8464	0.5138	0.0484	0.4217
51	25.8750	21.6288	0.9623	0.4926	0.9085	0.0716
52	16.4663	3.5816	0.8973	0.8633	0.3797	0.1826
53	15.8949	6.6661	0.7135	0.4223	0.4709	0.7934
54	18.4263	18.6462	0.9798	0.2435	0.9023	0.9066
55	10.4941	7.1066	0.5522	0.6016	0.1596	0.5752
56	12.8744	9.5374	0.0257	0.8938	0.1710	0.8881
57	11.9316	7.6043	0.6360	0.3746	0.1171	0.4774
58	5.3345	11.0907	0.4931	0.5313	0.6631	0.4787
59	8.0745	5.2730	0.2948	0.8758	0.2607	0.1512
60	10.8981	12.0806	0.0544	0.1278	0.6542	0.8694
61	19.0876	6.6191	0.9651	0.0933	0.4805	0.8090
62	8.3979	25.1441	0.5107	0.1299	0.9489	0.6049
63	10.4624	11.3307	0.5552	0.5946	0.6262	0.2679
64	7.2887	8.6390	0.2523	0.8471	0.5271	0.0787
65	6.6342	17.8156	0.4500	0.6056	0.7820	0.4937
66	10.4126	3.8363	0.1277	0.5998	0.2494	0.1996
67	5.9697	9.5520	0.4739	0.0363	0.6024	0.5735
68	6.6959	18.8322	0.5951	0.0162	0.8864	0.0864
69	8.4479	6.4422	0.1407	0.1824	0.1488	0.9888
70	6.0636	4.5297	0.2053	0.0736	0.3508	0.7970
71	23.8139	12.7176	0.9525	0.4358	0.7193	0.4946
72	5.9689	3.7303	0.5218	0.1567	0.0956	0.1070
73	6.7916	18.3957	0.5555	0.7630	0.8152	0.3555
74	11.1203	6.1022	0.0547	0.1246	0.5182	0.3672
75	4.9838	12.7269	0.3204	0.2748	0.6129	0.1217
76	5.9248	11.1394	0.2471	0.5091	0.6191	0.7990

Lanjutan Lampiran 11

No.	y_1	y_2	x_1	x_2	x_3	x_4
77	3.8641	10.0373	0.2967	0.4944	0.5924	0.0792
78	9.3686	23.3533	0.1660	0.7981	0.9158	0.3339
79	8.1343	14.1535	0.4470	0.4985	0.6284	0.0017
80	6.1016	7.4135	0.2309	0.6785	0.2971	0.9164
81	6.1482	9.7130	0.2150	0.2498	0.0973	0.8228
82	18.9421	6.2442	0.8481	0.2508	0.4931	0.4080
83	8.1767	5.4815	0.1826	0.0642	0.5207	0.4837
84	7.7219	22.0986	0.1867	0.2375	0.9117	0.8305
85	7.2051	16.3052	0.1572	0.1217	0.7652	0.5364
86	18.2804	5.8741	0.8637	0.3857	0.2205	0.4186
87	8.3472	18.7259	0.5383	0.8199	0.8023	0.9544
88	8.8602	18.2785	0.1018	0.1302	0.8156	0.8932
89	13.4103	4.2915	0.7590	0.8176	0.3428	0.1781
90	7.0867	22.6086	0.1996	0.7838	0.9151	0.3453
91	4.3797	4.8795	0.4518	0.1516	0.2084	0.4497
92	18.0798	10.5233	0.8579	0.7550	0.5961	0.5264
93	8.9944	10.0504	0.1042	0.8255	0.1501	0.7023
94	4.0784	7.0006	0.4258	0.7521	0.3588	0.9151
95	5.1178	4.3022	0.4618	0.7520	0.3166	0.3085
96	7.3995	6.7049	0.4858	0.9769	0.4292	0.4182
97	20.3674	3.5734	0.8350	0.5139	0.1383	0.1279
98	6.3739	3.9692	0.4189	0.5644	0.1197	0.2489
99	9.5887	6.3932	0.5797	0.8265	0.4715	0.5446
100	18.5305	24.3579	0.8750	0.7168	0.9399	0.0315
101	7.3181	8.6082	0.2969	0.5424	0.3480	0.5348
102	18.5195	7.6151	0.7867	0.4587	0.5225	0.7021
103	4.2772	11.6803	0.4070	0.9367	0.6378	0.8687
104	7.5494	7.3967	0.1848	0.2049	0.5373	0.3395
105	7.3281	8.5661	0.2416	0.2105	0.0937	0.9418
106	8.3759	1.3439	0.2348	0.8507	0.0450	0.0586
107	11.0257	16.8712	0.6924	0.9437	0.7475	0.2791
108	7.5997	12.7338	0.1289	0.0641	0.6645	0.1836
109	8.5578	6.3887	0.5130	0.5720	0.3780	0.9717
110	16.3205	6.9426	0.9445	0.9881	0.5047	0.3377
111	5.8280	4.3336	0.2205	0.2862	0.4750	0.3585
112	15.3186	5.3451	0.7866	0.6850	0.5165	0.3836
113	10.8391	15.0280	0.6367	0.5429	0.7650	0.6292
114	13.5864	8.1248	0.8585	0.0889	0.5789	0.9408
115	6.3605	12.9102	0.4590	0.7217	0.6275	0.2969

Lanjutan Lampiran 11

No.	y_1	y_2	x_1	x_2	x_3	x_4
116	5.8271	6.4790	0.4641	0.5189	0.3051	0.8082
117	5.1801	6.5072	0.3195	0.9098	0.3236	0.6114
118	13.4232	2.0181	0.8140	0.1375	0.2497	0.0395
119	7.4512	14.7321	0.1889	0.7853	0.7158	0.2561
120	6.5569	4.3591	0.3267	0.5909	0.3412	0.0170
121	22.6053	22.9046	0.9658	0.3058	0.9090	0.4346
122	10.8754	3.9775	0.0402	0.4527	0.3394	0.3627
123	12.1641	8.4500	0.0191	0.2188	0.3234	0.6797
124	11.0751	3.1145	0.0117	0.0797	0.4408	0.9250
125	4.2132	14.9934	0.3777	0.3805	0.7279	0.2434

Lampiran 12. Data Simulasi pada Fungsi Percobaan Trigonometri-Polinomial
 $(n_1 = n_2 = 125; \sigma_{11} = \sigma_{22} = 1; \rho = 0,5)$

No.	y_1	y_2	x_1	x_2	x_3	x_4
1	-1.9326	-2.3891	0.7639	0.7618	0.5890	0.2996
2	4.1416	1.7328	0.2501	0.4676	0.4180	0.9512
3	5.4300	5.7555	0.0802	0.5313	0.1466	0.5441
4	-0.1399	2.2158	0.4625	0.9571	0.0819	0.3584
5	2.0773	6.3246	0.1187	0.1920	0.2659	0.2002
6	-1.3912	4.7384	0.0353	0.0581	0.1606	0.2001
7	4.3012	0.7756	0.2716	0.1379	0.4288	0.5684
8	1.3307	-6.0580	0.1258	0.9115	0.9273	0.8188
9	-0.5396	4.5069	0.6049	0.3980	0.2857	0.7218
10	-3.6653	-0.7794	0.7230	0.5121	0.9007	0.1162
11	-1.6938	3.3855	0.8261	0.6944	0.0088	0.1890
12	6.3034	7.5649	0.2174	0.3566	0.3074	0.1258
13	6.7740	2.9453	0.2610	0.3743	0.1309	0.5934
14	1.4197	-2.8972	0.0883	0.8073	0.5542	0.7184
15	-4.3705	2.9977	0.8177	0.1791	0.3066	0.4640
16	-1.2288	-0.9195	0.0668	0.9293	0.5500	0.9927
17	-0.7195	2.1608	0.9288	0.9470	0.2509	0.5651
18	-1.6122	4.8158	0.6149	0.7594	0.3256	0.9004
19	4.6454	-1.4073	0.4709	0.7230	0.6023	0.2575
20	-2.6114	2.6206	0.4201	0.0167	0.2821	0.4840
21	-5.1040	-1.2705	0.8726	0.0818	0.7360	0.1773
22	4.4251	-6.1833	0.2417	0.2448	0.6488	0.8559
23	2.6744	-6.6309	0.4317	0.4638	0.8855	0.7027
24	-1.6953	-5.3205	0.6904	0.3762	0.7914	0.8558
25	6.0725	-4.6230	0.2710	0.7020	0.6509	0.3850
26	5.6298	6.8155	0.3221	0.7122	0.1734	0.1042
27	-4.9348	4.2751	0.8846	0.0924	0.1601	0.2889
28	3.6338	-1.6484	0.0904	0.6013	0.5000	0.7284
29	0.5881	5.9197	0.1677	0.0174	0.2658	0.2901
30	0.6286	-4.9410	0.8608	0.6791	0.7560	0.2952
31	-5.2761	4.6852	0.8994	0.0141	0.3228	0.7877
32	0.8836	3.9026	0.9231	0.9054	0.1600	0.5752
33	-3.6249	4.1632	0.8247	0.1943	0.4410	0.2054
34	4.5555	5.2208	0.1130	0.4770	0.0305	0.0661
35	-2.2076	1.5537	0.5381	0.1485	0.0486	0.3791
36	2.3968	3.4545	0.1346	0.4450	0.1536	0.4226
37	1.2078	-4.0634	0.1211	0.8025	0.9480	0.5192

Lanjutan Lampiran 12

No.	y_1	y_2	x_1	x_2	x_3	x_4
38	2.1552	0.1096	0.0996	0.1894	0.7823	0.0448
39	3.0677	1.7860	0.1290	0.6354	0.0323	0.8909
40	-2.9890	3.5984	0.6999	0.7427	0.2701	0.1861
41	5.0635	-4.8774	0.3567	0.5286	0.9689	0.6138
42	1.5524	-4.6427	0.4886	0.5097	0.6730	0.7890
43	2.5576	0.8636	0.2346	0.9315	0.4016	0.5046
44	-5.7454	-8.0291	0.8756	0.2575	0.9416	0.6711
45	-5.2863	1.7666	0.9544	0.0057	0.2943	0.5702
46	1.7500	-5.4942	0.4309	0.7202	0.6044	0.6570
47	2.7117	4.5059	0.0251	0.4990	0.3182	0.2004
48	-1.0605	-2.1284	0.0944	0.9026	0.6111	0.0983
49	-1.2755	-4.1839	0.7862	0.7426	0.6176	0.6679
50	-1.1955	1.9926	0.0497	0.0321	0.1278	0.6317
51	-5.2579	-4.6450	0.6685	0.1007	0.7766	0.5046
52	2.3661	-4.6976	0.2297	0.8151	0.7566	0.1972
53	-2.3712	4.5728	0.6986	0.8855	0.2141	0.8634
54	-2.8273	1.8734	0.6630	0.5728	0.9661	0.0756
55	-1.0721	1.5775	0.9420	0.2788	0.3556	0.6312
56	0.7836	-4.6141	0.0596	0.2375	0.5583	0.5836
57	-6.1132	-5.1448	0.6291	0.1360	0.6732	0.3664
58	-0.6301	2.7043	0.8932	0.5248	0.1570	0.5333
59	-4.9323	-4.4053	0.8189	0.2334	0.6228	0.8463
60	-1.2095	2.8818	0.5755	0.9776	0.0638	0.4145
61	6.3834	4.2328	0.2671	0.6965	0.2006	0.5838
62	-2.6465	-5.3278	0.8137	0.8717	0.8153	0.4651
63	4.8201	2.4762	0.3735	0.7963	0.4241	0.8559
64	-2.2487	-4.0377	0.6753	0.4707	0.9195	0.7449
65	-0.7022	-2.2718	0.8900	0.6147	0.5785	0.3885
66	-3.2082	1.8136	0.6568	0.7807	0.1053	0.6605
67	-4.4320	-1.1961	0.6858	0.2581	0.5378	0.2219
68	-2.7047	5.3038	0.7634	0.6811	0.2703	0.2749
69	3.3860	3.1421	0.2386	0.7532	0.1022	0.5779
70	-0.4348	-2.7932	0.9012	0.9061	0.6233	0.2911
71	2.9038	-6.4450	0.2458	0.1968	0.8779	0.8249
72	-2.9474	-8.5021	0.7390	0.7202	0.8080	0.6599
73	0.7419	-6.8678	0.0884	0.1583	0.7262	0.4276
74	5.2102	4.4323	0.1894	0.3341	0.3224	0.8863
75	-1.2752	-5.5620	0.4359	0.2239	0.9791	0.4984
76	-0.9287	-4.6956	0.6537	0.7237	0.5891	0.5793

Lanjutan Lampiran 12

No.	y_1	y_2	x_1	x_2	x_3	x_4
77	-2.1124	-2.3183	0.8589	0.5971	0.5041	0.5423
78	2.1974	1.7093	0.1000	0.3437	0.3499	0.8389
79	-0.0047	2.6554	0.0535	0.1343	0.2851	0.6317
80	-3.2235	-5.2718	0.6776	0.7779	0.7938	0.2818
81	6.2519	1.7287	0.3139	0.6329	0.0423	0.3338
82	-0.8766	-6.1303	0.5581	0.5307	0.7342	0.9130
83	-0.9969	0.0329	0.9939	0.1775	0.4111	0.4893
84	-1.1317	-4.8357	0.8678	0.4699	0.7946	0.8827
85	0.3636	-3.4935	0.0128	0.3654	0.5251	0.4139
86	-0.8278	0.3053	0.5119	0.9354	0.4686	0.2171
87	-2.5445	-3.4897	0.7751	0.2866	0.8098	0.2481
88	5.2726	-0.6533	0.1852	0.4752	0.5184	0.5022
89	6.2374	5.2445	0.2933	0.5418	0.1270	0.1449
90	-2.6118	-2.6669	0.6186	0.4539	0.5961	0.9457
91	-2.3442	-2.1158	0.6520	0.8205	0.5884	0.2320
92	-1.6373	-7.3254	0.5591	0.4506	0.7672	0.9838
93	3.4066	0.2296	0.4093	0.7908	0.9811	0.1721
94	4.3515	4.7615	0.1423	0.5378	0.2647	0.8986
95	-3.0683	-5.3778	0.7814	0.2907	0.6761	0.4404
96	4.8243	-4.2417	0.3155	0.8351	0.6002	0.3629
97	2.9172	-8.0933	0.2995	0.1880	0.8052	0.4801
98	1.7287	2.8350	0.4948	0.4290	0.1163	0.4607
99	-2.8934	-5.9697	0.8293	0.6781	0.9715	0.8542
100	-1.6948	-3.4547	0.0193	0.2227	0.5173	0.6766
101	3.2895	-6.4075	0.2945	0.9780	0.7255	0.9423
102	-4.2014	-0.0037	0.8768	0.1858	0.5784	0.1752
103	-5.0716	2.3633	0.8580	0.2878	0.2625	0.5749
104	3.9958	2.1561	0.2408	0.9409	0.4945	0.2073
105	4.3609	-3.2282	0.3664	0.6790	0.6991	0.2409
106	3.4781	4.6478	0.1412	0.7085	0.2945	0.9661
107	5.1892	-5.9483	0.3398	0.5933	0.8499	0.9698
108	2.5493	-2.7849	0.4519	0.7413	0.8655	0.1277
109	0.5844	0.6283	0.0443	0.8598	0.6113	0.1287
110	0.8029	-5.9906	0.8732	0.5120	0.8295	0.8863
111	-3.6126	2.8655	0.5990	0.8720	0.1823	0.2687
112	-0.6661	1.0433	0.7062	0.6569	0.5404	0.1504
113	3.7990	6.0685	0.4277	0.2402	0.3341	0.9162
114	-1.8221	0.3987	0.6892	0.9943	0.3762	0.6524
115	-0.8222	-3.4589	0.5089	0.3348	0.9298	0.3058

Lanjutan Lampiran 12

No.	y_1	y_2	x_1	x_2	x_3	x_4
116	-2.8050	-5.7940	0.8143	0.5566	0.7452	0.6248
117	1.5601	3.7774	0.9422	0.3946	0.9152	0.0288
118	0.7377	4.7656	0.3648	0.0502	0.0258	0.0358
119	-6.1864	-8.2701	0.7185	0.1954	0.8248	0.6481
120	-3.6817	3.7706	0.9793	0.0561	0.0369	0.0606
121	6.7041	-3.0751	0.2561	0.4226	0.9997	0.9300
122	-7.0153	3.7532	0.6659	0.0581	0.2332	0.4966
123	2.1291	-2.2039	0.0638	0.4941	0.5910	0.1627
124	3.6244	5.3279	0.4687	0.4193	0.0400	0.9638
125	1.1526	-1.7849	0.2234	0.0595	0.6177	0.0715

Lampiran 13. Data Simulasi pada Fungsi Percobaan Eksponensial-Trigonometri
 $(n_1 = n_2 = 50; \sigma_{11} = \sigma_{11} = 1; \rho = 0,5)$

No.	y_1	y_2	x_1	x_2	x_3	x_4
1	4.2632	-4.8938	0.2739	0.3286	0.5529	0.1873
2	6.3259	4.1579	0.1321	0.9275	0.1423	0.9436
3	8.2931	2.8706	0.6997	0.7561	0.3804	0.9479
4	-6.1577	-3.0069	0.4859	0.2882	0.3966	0.4530
5	6.4757	-0.0916	0.1827	0.6062	0.5767	0.8108
6	5.4802	1.1443	0.1012	0.7660	0.0194	0.9289
7	5.4997	-1.2307	0.2016	0.8462	0.5776	0.6727
8	7.6614	5.6440	0.1347	0.9020	0.9322	0.3723
9	2.6472	4.6717	0.3238	0.5957	0.1069	0.4057
10	6.3143	5.9098	0.9505	0.0685	0.7321	0.4388
11	-3.8057	8.1442	0.5321	0.2180	0.9705	0.6786
12	2.2926	-3.1497	0.2477	0.8694	0.6089	0.4651
13	-4.6088	7.2204	0.4373	0.4142	0.7197	0.9533
14	3.5533	3.2426	0.6691	0.6612	0.3028	0.3547
15	-3.2126	-3.5624	0.5477	0.7835	0.4590	0.3390
16	-1.4377	1.7287	0.6091	0.2479	0.0480	0.8959
17	9.0689	-0.0445	0.8631	0.5544	0.3854	0.5454
18	-2.1459	-1.7720	0.3807	0.2296	0.3617	0.7493
19	7.1240	1.2480	0.7490	0.0069	0.2876	0.1249
20	7.3304	9.4610	0.1567	0.7666	0.8167	0.4532
21	6.7417	-4.2935	0.0581	0.0218	0.4505	0.0747
22	-0.2321	9.3518	0.3397	0.3931	0.8066	0.6633
23	7.1124	8.6665	0.8172	0.2525	0.7902	0.7037
24	-2.1320	3.6750	0.3775	0.2042	0.2830	0.9190
25	6.8966	4.3784	0.9726	0.6623	0.0683	0.6601
26	2.1703	5.5627	0.6053	0.9147	0.0549	0.6901
27	3.6142	3.2912	0.3382	0.0069	0.6375	0.8537
28	9.9602	-0.3242	0.9280	0.7464	0.4243	0.4679
29	10.1866	6.8143	0.8984	0.7997	0.9055	0.4585
30	10.5248	-1.6425	0.8507	0.9078	0.4173	0.8061
31	6.3946	6.8819	0.2568	0.9746	0.1541	0.8248
32	5.6004	-4.1801	0.2855	0.1199	0.5400	0.1904
33	8.6239	2.4267	0.7799	0.5190	0.9371	0.0257
34	6.8957	-2.2313	0.7014	0.8220	0.6610	0.0568
35	-5.3227	-5.7859	0.4925	0.6370	0.3947	0.1429
36	8.8333	3.2192	0.9677	0.9539	0.2590	0.1714
37	-2.8463	9.3926	0.4762	0.9469	0.8479	0.6258

Lanjutan Lampiran 13

No.	y_1	y_2	x_1	x_2	x_3	x_4
38	7.5736	2.5000	0.9949	0.9666	0.9451	0.0295
39	-2.8319	0.1207	0.4906	0.0673	0.3770	0.4723
40	-5.0358	3.8255	0.5035	0.4375	0.0673	0.6784
41	8.3228	5.2803	0.7688	0.3208	0.1816	0.1148
42	-2.6420	-2.8776	0.3881	0.1341	0.5757	0.2361
43	-4.0659	4.3315	0.4533	0.1346	0.1859	0.2891
44	6.7565	1.9109	0.1329	0.8059	0.2914	0.1728
45	8.4485	-3.6160	0.7585	0.5248	0.4617	0.3237
46	-1.4876	1.2124	0.5652	0.9443	0.3470	0.8011
47	2.5374	2.8140	0.6486	0.9883	0.3182	0.2996
48	9.4529	-2.3054	0.7981	0.4099	0.4599	0.7756
49	6.2873	5.2030	0.2204	0.3712	0.2359	0.5528
50	5.6781	0.1577	0.8579	0.2269	0.0278	0.5547
51	7.8910	3.9449	0.9047	0.4460	0.6585	0.7306
52	1.3945	6.6672	0.2920	0.2662	0.1588	0.7736
53	6.3486	11.7612	0.7259	0.4591	0.8027	0.9008
54	2.1313	-3.3420	0.3394	0.4329	0.4086	0.1382
55	3.8488	1.0231	0.2727	0.2596	0.3274	0.7941
56	6.1480	4.3714	0.1703	0.1337	0.7460	0.1894
57	3.0948	1.7530	0.6640	0.4192	0.7464	0.0290
58	-4.2389	3.5689	0.5359	0.5069	0.1740	0.1274
59	7.2211	3.9071	0.8291	0.3243	0.1175	0.1337
60	3.5208	2.5007	0.2674	0.6847	0.1740	0.1283
61	5.1104	3.4991	0.1762	0.4431	0.6274	0.9353
62	-3.1315	6.8262	0.4312	0.4357	0.8419	0.2732
63	-3.5317	-1.5003	0.4757	0.7930	0.5101	0.9427
64	10.6668	6.9278	0.7852	0.8156	0.1658	0.6382
65	6.2923	7.8243	0.1307	0.7521	0.7143	0.8725
66	2.8685	6.0783	0.0514	0.7893	0.9070	0.3671
67	-1.6512	4.6189	0.6275	0.5013	0.2185	0.2362
68	0.9822	5.4775	0.0291	0.5552	0.8710	0.1873
69	6.4703	6.4579	0.1362	0.6307	0.2118	0.5456
70	2.8492	6.4594	0.6946	0.0980	0.8367	0.2551
71	-3.4440	7.2522	0.5157	0.2457	0.8593	0.3058
72	-5.0044	-5.9814	0.5426	0.6157	0.5234	0.0155
73	7.5737	-3.4301	0.8085	0.3050	0.4774	0.5875
74	9.9647	11.7637	0.7937	0.7670	0.8899	0.9626
75	-5.5119	4.7084	0.5019	0.2672	0.0651	0.8498
76	5.3468	-7.1267	0.2766	0.0395	0.5095	0.0079

Lanjutan Lampiran 13

No.	y_1	y_2	x_1	x_2	x_3	x_4
77	7.5385	2.4217	0.1197	0.2966	0.6208	0.6340
78	7.6823	4.6950	0.8866	0.5564	0.7336	0.3593
79	7.8898	3.0653	0.9703	0.9691	0.2300	0.1141
80	6.2271	0.3790	0.9425	0.6891	0.0219	0.5408
81	2.3449	6.1214	0.6381	0.7179	0.1390	0.4164
82	3.6907	8.6871	0.0906	0.5590	0.7695	0.5171
83	3.3295	8.6171	0.0747	0.5334	0.9698	0.8861
84	8.4763	-2.0570	0.1825	0.8757	0.3868	0.1494
85	2.7632	2.4052	0.0317	0.3931	0.9934	0.4347
86	4.7490	-0.9285	0.7249	0.4581	0.3264	0.0590
87	7.5526	5.8654	0.1442	0.2082	0.1372	0.3810
88	1.0011	-1.1271	0.6359	0.7573	0.3848	0.7224
89	8.6536	-3.7330	0.7898	0.5467	0.5626	0.0951
90	-1.4966	3.1066	0.5663	0.3574	0.6338	0.6672
91	1.6930	-1.8861	0.3774	0.7010	0.5416	0.2964
92	7.2708	1.7090	0.8216	0.1092	0.3150	0.5986
93	4.4006	3.9808	0.3049	0.0066	0.1593	0.1519
94	3.1755	5.5428	0.3194	0.5973	0.1526	0.4364
95	9.6252	2.3540	0.7850	0.6592	0.1370	0.0127
96	-4.1679	2.3885	0.5037	0.5800	0.7098	0.2290
97	5.2348	-3.4656	0.2610	0.9100	0.4649	0.2637
98	5.9741	2.4649	0.7325	0.6360	0.1133	0.5114
99	4.3503	-0.2380	0.1629	0.5256	0.7009	0.2151
100	5.8700	5.2783	0.9211	0.2596	0.1800	0.3461
101	5.9707	9.2943	0.2222	0.0512	0.8037	0.7478
102	4.4582	-4.2649	0.0836	0.7320	0.5140	0.4136
103	7.0775	-3.8250	0.0737	0.1643	0.5484	0.0558
104	6.7331	6.2957	0.7696	0.2804	0.2078	0.3900
105	9.2197	8.7618	0.8177	0.2594	0.7846	0.4745
106	5.5897	-3.2808	0.7404	0.5471	0.5265	0.8253
107	8.7308	-2.7051	0.7582	0.5413	0.5710	0.3036
108	5.7774	-2.5144	0.9612	0.7881	0.4220	0.8218
109	-3.1416	5.2649	0.4664	0.8696	0.7212	0.5657
110	10.0687	1.4814	0.7870	0.7875	0.0731	0.0544
111	-0.9838	-2.5625	0.4226	0.9694	0.5949	0.2600
112	4.6941	10.4552	0.9437	0.1805	0.8620	0.5891
113	-0.9986	-3.5792	0.0013	0.9306	0.4488	0.4797
114	2.1444	-0.7739	0.9813	0.0452	0.6526	0.1987
115	-4.1519	1.5603	0.5702	0.2406	0.3035	0.2390

Lanjutan Lampiran 13

No.	y_1	y_2	x_1	x_2	x_3	x_4
116	1.0315	-0.8706	0.3465	0.0089	0.6074	0.7802
117	-3.0913	2.5388	0.5575	0.6716	0.2789	0.6173
118	2.8148	4.9711	0.2998	0.9048	0.7996	0.1441
119	7.8815	10.1765	0.1591	0.5724	0.7962	0.7161
120	1.6545	4.2102	0.6653	0.1555	0.9541	0.4015
121	3.7093	-1.6169	0.6842	0.5024	0.4443	0.4624
122	7.8875	-2.4919	0.7924	0.5677	0.4569	0.7073
123	0.4257	-0.8403	0.3486	0.1883	0.5998	0.4012
124	5.2923	3.5514	0.2501	0.3242	0.8426	0.0144
125	0.1394	-0.5773	0.3450	0.7160	0.0312	0.0746

Lampiran 14. Data Simulasi pada Fungsi Percobaan Polinomial-Eksponensial
 $(n_1 = n_2 = 50; \sigma_{11} = \sigma_{11} = 1; \rho = 0,5)$

No.	y_1	y_2	x_1	x_2	x_3	x_4
1	3.8789	2.3777	0.3487	0.0191	0.4910	0.2856
2	4.0067	3.8836	0.1268	0.1584	0.8205	0.5481
3	2.8687	7.8296	0.5579	0.9043	0.0734	0.8868
4	3.2628	5.7483	0.4289	0.1634	0.8980	0.5690
5	2.4313	2.7815	0.4267	0.5730	0.7135	0.0503
6	5.4577	11.6638	0.7802	0.6843	0.0452	0.0138
7	10.1963	6.9679	0.9913	0.7049	0.1224	0.3193
8	5.9179	4.3831	0.8850	0.1808	0.7170	0.7513
9	3.1871	4.7041	0.1787	0.2707	0.9692	0.1069
10	2.1356	3.3672	0.5309	0.4739	0.4568	0.3964
11	-0.4062	2.8496	0.2654	0.5864	0.3446	0.1859
12	6.1270	1.9821	0.9820	0.3334	0.6985	0.1706
13	3.6748	6.0247	0.1354	0.7237	0.9035	0.7241
14	0.1740	1.0626	0.5408	0.1186	0.5067	0.4725
15	2.2307	7.3722	0.4692	0.7764	0.0631	0.3011
16	7.1159	2.9489	0.8048	0.9881	0.7815	0.1974
17	7.1130	2.6145	0.8774	0.7572	0.3472	0.6069
18	4.0392	5.8362	0.7159	0.2114	0.2025	0.6112
19	3.5019	2.4837	0.1208	0.7368	0.6269	0.7061
20	5.7701	2.4246	0.9270	0.1835	0.9145	0.2711
21	4.9922	7.0479	0.8135	0.4633	0.1753	0.2904
22	6.7781	2.4593	0.9822	0.3784	0.7652	0.2793
23	3.0532	5.3487	0.7644	0.2398	0.2183	0.9372
24	1.4688	4.3509	0.4812	0.5657	0.5265	0.6909
25	5.4807	5.1116	0.1332	0.0715	0.8791	0.6588
26	0.4041	2.1546	0.4208	0.1734	0.3437	0.3710
27	2.8839	6.2419	0.0624	0.9291	0.9030	0.8962
28	2.5989	5.3516	0.2736	0.1292	0.7619	0.6051
29	0.5176	1.7090	0.3021	0.4552	0.4473	0.4306
30	3.4931	3.7979	0.0705	0.2075	0.4228	0.6660
31	5.9165	6.5653	0.6929	0.9588	0.2409	0.0774
32	2.7802	3.2593	0.1058	0.5724	0.2402	0.5382
33	2.2737	4.2271	0.0833	0.8192	0.2341	0.1251
34	1.7241	4.0522	0.2889	0.2406	0.9613	0.5966
35	3.3483	1.9207	0.8433	0.3405	0.7429	0.2083
36	2.6344	3.4211	0.3826	0.0215	0.5545	0.0105
37	0.5552	1.8774	0.4584	0.3388	0.7411	0.0552

Lanjutan Lampiran 14

No.	y_1	y_2	x_1	x_2	x_3	x_4
38	3.7191	7.7014	0.7240	0.1199	0.0217	0.8633
39	2.9525	2.3466	0.1389	0.4025	0.3915	0.3331
40	1.0179	5.5662	0.7739	0.2121	0.1369	0.5955
41	1.9468	4.8838	0.4727	0.6591	0.3358	0.9856
42	5.4165	4.8205	0.9746	0.4582	0.1048	0.6164
43	8.8954	4.6075	0.9013	0.8777	0.1728	0.3589
44	4.7311	4.9244	0.6816	0.6350	0.2505	0.0863
45	3.3959	4.3365	0.1142	0.9928	0.5148	0.8668
46	7.3440	4.4863	0.8206	0.8697	0.4111	0.9473
47	0.4156	2.7561	0.4512	0.1961	0.4680	0.6911
48	4.4164	3.8469	0.4009	0.9650	0.5157	0.6006
49	3.4018	2.8239	0.7956	0.2703	0.7249	0.1268
50	1.9768	4.0684	0.6337	0.7170	0.2267	0.3698
51	3.7757	4.1326	0.6636	0.5796	0.7176	0.9739
52	0.9497	5.9190	0.5942	0.3509	0.1020	0.6396
53	5.7386	6.8695	0.7600	0.7151	0.0855	0.3845
54	6.2308	2.3703	0.5791	0.9359	0.5062	0.2303
55	4.6422	4.3016	0.8412	0.2988	0.4025	0.6193
56	3.7256	2.3350	0.1159	0.5356	0.7299	0.2542
57	6.4376	7.1894	0.9984	0.1253	0.9810	0.8617
58	5.3761	6.7242	0.7782	0.7740	0.9706	0.6167
59	0.0512	2.0404	0.3282	0.0903	0.3125	0.7539
60	6.4857	3.3117	0.0125	0.0348	0.6653	0.3929
61	0.5536	3.7308	0.6288	0.0853	0.4129	0.7391
62	8.6845	5.4165	0.8817	0.8593	0.3000	0.3134
63	2.8219	1.5471	0.5686	0.0819	0.5313	0.3357
64	4.1622	7.6618	0.0978	0.8147	0.1014	0.6297
65	10.7344	7.6253	0.9986	0.8487	0.0646	0.8565
66	4.1933	4.3475	0.7573	0.1203	0.9598	0.0525
67	2.6945	7.1080	0.4047	0.7867	0.0113	0.6556
68	5.2271	5.2334	0.8665	0.4157	0.9086	0.5293
69	2.7697	3.2385	0.3938	0.8294	0.2519	0.2881
70	2.2300	8.5989	0.5921	0.5805	0.0139	0.7621
71	6.5895	-0.0626	0.9420	0.0067	0.4772	0.1756
72	5.2505	6.0080	0.8596	0.4627	0.2203	0.9958
73	6.5928	2.3868	0.8788	0.5501	0.6322	0.5978
74	5.4684	2.3966	0.7598	0.8558	0.7788	0.7658
75	3.6806	4.8609	0.6032	0.9162	0.2703	0.5969
76	7.6393	1.1165	0.8288	0.7694	0.6328	0.2980

Lanjutan Lampiran 14

No.	y_1	y_2	x_1	x_2	x_3	x_4
77	2.8103	4.8498	0.3090	0.4581	0.2116	0.9199
78	2.1981	3.0851	0.1370	0.6378	0.6574	0.6248
79	0.9784	3.7413	0.4046	0.9217	0.3368	0.2788
80	3.6377	6.2907	0.5801	0.6684	0.1907	0.5554
81	2.1717	3.3664	0.2637	0.6694	0.9070	0.5207
82	0.4401	6.4998	0.3502	0.3676	0.1619	0.6284
83	9.1309	3.0253	0.9905	0.5230	0.4638	0.6120
84	2.5612	1.9758	0.3803	0.8308	0.7416	0.3521
85	1.8669	5.2557	0.6060	0.6308	0.2505	0.1645
86	1.4688	6.6719	0.1286	0.2838	0.9575	0.7682
87	5.9633	6.4831	0.8796	0.4970	0.1370	0.1861
88	2.0837	4.9911	0.6225	0.5539	0.1173	0.9318
89	5.3533	2.5208	0.6390	0.8353	0.6496	0.1260
90	1.6101	1.7767	0.5829	0.5543	0.7963	0.5440
91	1.5947	4.0537	0.2252	0.2534	0.6966	0.9277
92	3.0479	2.5047	0.0570	0.0370	0.2105	0.2871
93	5.7218	1.7528	0.8002	0.9441	0.6137	0.4410
94	2.2607	7.6334	0.7057	0.2325	0.9852	0.9332
95	3.2867	1.0819	0.3561	0.8063	0.8008	0.2559
96	1.1150	1.2229	0.5229	0.5595	0.4310	0.4331
97	1.3475	1.7731	0.3625	0.2911	0.6072	0.2816
98	0.4375	2.9083	0.2735	0.2365	0.6501	0.7458
99	7.0650	5.1064	0.8667	0.7220	0.3218	0.8538
100	4.5721	4.0419	0.0005	0.9279	0.8069	0.8279
101	9.0280	4.3268	0.9385	0.9183	0.2689	0.1153
102	4.3513	6.2883	0.5768	0.7954	0.2224	0.8455
103	2.3285	2.0016	0.7110	0.2335	0.5708	0.6398
104	2.9423	3.6966	0.6608	0.5507	0.8222	0.9470
105	4.4012	9.3470	0.2550	0.8821	0.0034	0.3027
106	3.0850	5.1807	0.2852	0.5527	0.1715	0.8829
107	1.6551	4.0534	0.5527	0.1602	0.9853	0.4583
108	3.6112	7.4884	0.7191	0.0408	0.0212	0.8869
109	2.3731	1.5480	0.1870	0.0362	0.9222	0.1923
110	3.3791	4.6277	0.2806	0.7898	0.5396	0.9555
111	-0.6393	2.6175	0.5613	0.1274	0.4412	0.9370
112	5.9352	4.9931	0.7467	0.8608	0.3465	0.1400
113	4.7614	3.5850	0.7100	0.6908	0.9144	0.5906
114	2.2899	3.3504	0.0698	0.9294	0.3401	0.8364
115	3.8682	1.7059	0.5536	0.8174	0.8313	0.3068

Lanjutan Lampiran 14

No.	y_1	y_2	x_1	x_2	x_3	x_4
116	0.7872	2.4606	0.3503	0.1731	0.4239	0.9212
117	2.1154	3.5800	0.3053	0.4857	0.4921	0.6044
118	5.3032	4.1207	0.1153	0.7503	0.6326	0.5768
119	4.2450	3.5524	0.4108	0.9483	0.5792	0.6872
120	1.8166	6.3413	0.2423	0.4697	0.1344	0.8213
121	3.1659	8.1606	0.5176	0.8405	0.0186	0.1007
122	3.5729	5.1791	0.6685	0.4838	0.9912	0.2065
123	1.1790	1.6090	0.4627	0.4924	0.6715	0.5155
124	-0.2147	1.2415	0.5555	0.3280	0.4995	0.3325
125	1.5745	2.9680	0.6079	0.1812	0.2218	0.2325

Lampiran 15. Data Persentase Penduduk Miskin (y_1), Indeks Pembangunan Manusia (y_2), Pertumbuhan Ekonomi (x_1), Angka Beban Ketergantungan (x_2), Persentase Pengeluaran Perkapita untuk Makanan (x_3), Tingkat Pengangguran Terbuka (x_4) dan Persentase Penduduk yang Pekerjaan Utamanya di Sektor Pertanian (x_5) menurut Kabupaten/Kota di Pulau Jawa

Kode	Kabupaten/Kota	y_1	y_2	x_1	x_2	x_3	x_4	x_5
3101	KEPULAUAN SERIBU	11.62	71.45	0.1	53.34	58.16	13.08	26.84
3171	Kota JAKARTA SELATAN	3.49	80.17	6.69	37.02	44.84	8.77	0.11
3172	Kota JAKARTA TIMUR	3.12	79.8	6.56	38.83	45.06	9.95	0.96
3173	Kota JAKARTA PUSAT	3.72	79.12	6.76	36.49	48.11	10.5	0.00
3174	Kota JAKARTA BARAT	3.47	79.43	6.39	35.66	47.04	8.95	0.62
3175	Kota JAKARTA UTARA	5.14	78.25	6.01	36.21	50.81	9.98	0.40
3201	BOGOR	8.82	73.08	5.99	53.74	56.1	9.12	13.64
3202	SUKABUMI	9.78	71.5	4.34	57.02	61.23	9.51	34.48
3203	CIANJUR	13.17	70.02	5.08	57.2	63.2	10.73	45.59
3204	BANDUNG	8.32	74.73	6.15	53.98	57.62	11.61	15.53
3205	GARUT	12.7	72.12	4.61	63.16	62.8	6.22	38.99
3206	TASIKMALAYA	11.75	72.84	4.32	57.41	64.06	4.66	30.55
3207	CIAMIS	9.61	72.14	4.99	52.99	62.34	4.97	35.18
3208	KUNINGAN	13.69	71.99	4.73	54.87	62.6	6.81	35.7
3209	CIREBON	14.94	69.58	4.81	51.29	61.84	15.92	25.63
3210	MAJALENGKA	14.44	71.16	4.76	52.22	61.67	6.42	36.56
3211	SUMEDANG	11.85	72.95	4.69	51.64	61	7.14	37.15
3212	INDRAMAYU	15.42	68.89	5.03	48.73	64.26	7.6	38.98
3213	SUBANG	12.47	71.79	4.52	50.3	59.41	7.59	45.41
3214	PURWAKARTA	9.56	72.21	6.31	53.47	58.79	9.19	24.21
3215	KARAWANG	11.1	70.89	5.44	48.67	58.2	11.23	18.3
3216	BEKASI	5.25	74.13	6.22	46.68	56.09	7.92	10.58
3217	BANDUNG BARAT	13.33	74.03	6.04	56.15	61.12	9.99	26.42
3271	Kota BOGOR	8.47	76.47	6.15	45.98	54.9	9.5	1.56
3272	Kota SUKABUMI	8.41	75.73	5.29	51.27	58.09	11.85	3.85
3273	Kota BANDUNG	4.55	76.86	8.98	41.48	50.54	9.27	0.99
3274	Kota CIREBON	11.08	76.02	5.57	47.08	53.03	12.71	0.69
3275	Kota BEKASI	5.55	77.17	6.85	40.67	49.62	8.88	0.41
3276	Kota DEPOK	2.46	79.71	7.15	43.84	49.7	9.55	2.02
3277	Kota CIMAHI	6.67	76.28	5.24	44.15	51	8.67	1.85
3278	Kota TASIKMALAYA	18.82	75.35	5.89	51.06	57.78	6.12	7.75
3279	Kota BANJAR	7.78	72.1	5.26	52.34	60.35	6.2	13.86
3301	CILACAP	15.92	72.77	3.03	54.93	57.75	7.26	34.09

Lanjutan Lampiran 15

Kode	Kabupaten/Kota	y_1	y_2	x_1	x_2	x_3	x_4	x_5
3302	BANYUMAS	19.44	73.33	5.88	52.44	56.83	5.09	23.42
3303	PURBALINGGA	21.19	72.97	6.26	55.14	58.74	4.99	25.8
3304	BANJARNEGARA	18.87	70.7	5.25	52.31	59.53	3.66	49.42
3305	KEBUMEN	22.4	71.86	5.47	59.06	59.68	3.55	36.08
3306	PURWOREJO	16.32	73.53	5.04	56.53	59.36	3.17	39.49
3307	WONOSOBO	22.5	71.45	5.14	54.27	57.75	5.18	54.57
3308	MAGELANG	13.97	73.14	5.84	51.52	58.66	4.35	40.94
3309	BOYOLALI	13.88	71.5	5.66	53.41	55.58	4.41	27.72
3310	KLATEN	16.71	74.46	5.54	50.47	56.81	3.69	20.43
3311	SUKOHARJO	10.16	74.21	5.03	46.18	55.02	6.07	15.57
3312	WONOGIRI	14.67	72.59	5.89	53.28	57.48	3.43	52.97
3313	KARANGANYAR	14.07	74.62	5.82	49.03	54.28	5.8	28.2
3314	SRAGEN	16.72	71.85	6.6	50.95	57.1	5.85	37.06
3315	GROBOGAN	16.14	71.77	6.16	50.94	61.41	4.18	53.01
3316	BLORA	15.11	71.49	5.02	48.96	61.76	4.72	46.86
3317	REMBANG	21.88	72.81	4.88	45.07	59.11	5.73	43.5
3318	PATI	13.61	73.81	5.92	47.9	59.55	11.92	33.28
3319	KUDUS	8.63	73.69	4.33	43.47	55.14	5.88	12.3
3320	JEPARA	9.38	73.54	5.79	49.67	57.39	4.29	10.98
3321	DEMAK	16.73	73.52	4.64	49.1	59.8	8.39	36.3
3322	SEMARANG	9.4	74.98	6.02	47.63	54.59	4.86	37.75
3323	TEMANGGUNG	12.32	74.74	5.04	48.98	57.15	3.38	36.75
3324	KENDAL	13.17	71.48	5.54	48.14	56.91	6.29	43.13
3325	BATANG	12.4	71.41	5.02	47.9	63.34	5.87	30.15
3326	PEKALONGAN	13.86	72.37	5.32	52.73	61.37	5.06	18.51
3327	PEMALANG	19.28	70.66	5.28	54.7	63.52	4.82	40.06
3328	TEGAL	10.75	71.74	5.25	54.04	61.51	6.1	23.79
3329	BREBES	21.13	69.37	5.21	52.41	61.49	8.19	44.32
3371	Kota MAGELANG	10.31	77.26	6.51	45.11	50.8	8.96	0.18
3372	Kota SURAKARTA	12.01	78.6	6.12	39.83	46.97	6.27	0.73
3373	Kota SALATIGA	7.11	77.13	5.73	42.61	51.64	6.83	6.6
3374	Kota SEMARANG	5.13	77.98	6.42	39.38	50.95	6.00	3.66
3375	Kota PEKALONGAN	9.47	75.25	5.6	45.45	58.88	7.67	4.73
3376	Kota TEGAL	10.04	74.63	5.07	45.88	53.91	8.74	7.04
3401	KULON PROGO	23.32	75.33	5.01	54.41	57.45	3.04	45.76
3402	BANTUL	16.97	75.51	5.34	46.6	55.94	3.7	15.35
3403	GUNUNG KIDUL	22.72	71.11	4.84	54.52	62.34	1.38	55.96

Lanjutan Lampiran 15

Kode	Kabupaten/Kota	y_1	y_2	x_1	x_2	x_3	x_4	x_5
3404	SLEMAN	10.44	79.39	5.45	41.68	50.67	5.64	19.61
3471	Kota YOGYAKARTA	9.38	80.24	5.76	35.6	46.98	5.33	0.35
3501	PACITAN	17.22	72.78	6.73	51.41	61.77	1.02	59.53
3502	PONOROGO	11.72	71.91	6.52	48.84	58.07	3.14	52.62
3503	TRENGGALEK	14.15	74.09	6.62	47.79	59.64	2.98	56.69
3504	TULUNGAGUNG	9.37	74.45	6.99	50.41	55.18	3.1	41.62
3505	BLITAR	10.7	74.43	6.35	51.75	57.27	2.82	45.78
3506	KEDIRI	13.66	72.72	6.98	49.87	56.45	4.08	35.21
3507	MALANG	11	71.94	7.44	48.33	57.42	3.75	42.71
3508	LUMAJANG	12.36	69	6.43	46.26	62.77	4.6	50.57
3509	JEMBER	11.76	65.99	7.21	48.52	61.45	3.77	47.44
3510	BANYUWANGI	9.93	70.53	7.21	49.03	56.93	3.41	35.77
3511	BONDOWOSO	15.75	64.98	6.45	46.25	62.48	3.6	47.03
3512	SITUBONDO	14.29	65.06	6.54	42.95	61.31	3.33	46.31
3513	PROBOLINGGO	22.14	64.35	6.55	46.8	60.27	1.92	51.41
3514	PASURUAN	11.53	69.17	7.23	43.71	61.84	6.38	31.44
3515	SIDOARJO	6.42	77.36	7.13	42.27	54.72	5.37	5.35
3516	MOJOKERTO	10.67	74.42	7.21	44.8	59.24	3.35	22.19
3517	JOMBANG	12.18	73.86	6.97	48.74	58.11	6.72	27.61
3518	NGANJUK	13.17	71.96	6.68	49.1	57.68	4.09	46.4
3519	MADIUN	13.65	70.88	6.43	49	58.96	3.99	45.09
3520	MAGETAN	11.45	73.85	6.39	50.67	56.63	3.64	43.1
3521	NGAWI	15.94	70.2	6.58	47.46	61.07	2.94	59.51
3522	BOJONEGORO	16.6	67.74	5.68	45.61	59.68	3.42	57.44
3523	TUBAN	17.77	69.18	6.22	44.44	63.55	4.13	51.74
3524	LAMONGAN	16.64	71.05	7.13	46.18	57.74	4.75	49.92
3525	GRESIK	14.29	75.97	7.43	44.93	54.53	6.78	17.36
3526	BANGKALAN	24.61	65.69	6.31	56.31	65.27	5.13	65.17
3527	SAMPANG	27.87	61.67	6.12	54.39	63.15	1.71	64.79
3528	PAMEKASAN	19.53	66.51	6.32	47.41	62.58	2.29	76.08
3529	SUMENEP	21.87	66.41	6.33	42	62.04	1.14	79.16
3571	Kota KEDIRI	8.11	77.2	7.51	43.91	55.03	8.12	2.8
3572	Kota BLITAR	6.72	78.31	6.78	47.72	51.9	3.68	10.68
3573	Kota MALANG	5.19	78.43	7.57	39.19	50.36	7.96	1.25
3574	Kota PROBOLINGGO	18.33	75.44	6.89	46.48	53.77	5.26	6.42
3575	Kota PASURUAN	7.87	74.33	6.46	46.62	55.28	5.28	5.33
3576	Kota MOJOKERTO	6.46	78.01	7.11	44.32	52.62	7.52	2.37

Lanjutan Lampiran 15

Kode	Kabupaten/Kota	y_1	y_2	x_1	x_2	x_3	x_4	x_5
3577	Kota MADIUN	5.35	77.5	7.79	44.66	53.6	6.89	3.88
3578	Kota SURABAYA	6.23	78.33	7.62	38.24	50.19	5.27	0.42
3579	Kota BATU	4.45	75.42	8.25	46.01	55.56	3.51	27.01
3601	PANDEGLANG	9.27	69.22	5.62	61.45	63.88	8.87	36.59
3602	LEBAK	8.62	68.43	5.01	58.07	63.37	8.64	37.55
3603	TANGERANG	5.71	72.36	6.22	47.07	57.53	11.14	6.64
3604	SERANG	5.28	69.83	5.1	54.43	61.67	12.68	22.55
3671	Kota TANGERANG	5.55	75.72	6.41	39.8	52.72	8.13	0.67
3672	Kota CILEGON	3.81	75.89	6.82	46.71	60.7	11.05	4.62
3673	Kota SERANG	5.69	72.3	7.06	52.26	57.57	10.58	5.21
3674	Kota TANGERANG SELATAN	1.33	76.61	8.24	40.78	46.84	7.95	0.31

Lampiran 16. GUI MARS Birespon Kontinu

Untuk memudahkan pemrosesan data dalam pemodelan MARS birespon kontinu, maka dibuat fasilitas GUI (*Graphical User Interface*) yang komunikatif dengan pengguna (*user*). Tampilan GUI disajikan dalam gambar dibawah ini:

Pemodelan Regresi Nonparametrik Birespon Kontinu Menggunakan Metode MARS
(Oleh: Ampulembang, A.P., Otok, B.W., Rumiaty, A.T. & Budiasih)

Input Data

Data Respon:

☐ Prediktor Sama ☐ Prediktor Berbeda

Data Prediktor:

Input Kontrol

Max. Fungsi Basis:
Max. Interaksi:
Min. Observasi:

Output MARS Birespon Kontinu (MARS-BK)

Nilai RMSE: Direktori Output:
Nilai R2-Adj: Direktori Model:

Output Regresi Parametrik

	Linier	Kuadratik	Kubik
Nilai RMSE	<input type="text"/>	<input type="text"/>	<input type="text"/>
Nilai R2-Adj	<input type="text"/>	<input type="text"/>	<input type="text"/>

Tahapan penggunaan aplikasi GUI dari gambar diatas adalah sebagai berikut:

1. Input Data

- Tombol “Browse” pada data respon, digunakan untuk memasukkan data respon berupa file Microsoft Excel 2007. Setelah memilih file yang akan digunakan, maka nama file tersebut akan muncul dalam kotak dialog “Masukkan data respon”.

Lanjutan Lampiran 16.

- b. Kotak Pilihan “Prediktor sama” atau “Prediktor berbeda”, digunakan untuk menentukan apakah data prediktor yang digunakan untuk setiap respon adalah sama atau berbeda.
- c. Tombol “Browse” pada data prediktor, digunakan untuk memasukkan data prediktor berupa file Microsoft Excel 2007. Setelah memilih file yang akan digunakan, maka nama file tersebut akan muncul dalam kotak dialog “Masukkan data prediktor”.

2. Input Kontrol

- a. Max. Fungsi Basis, digunakan untuk memasukkan jumlah maksimum fungsi basis yang akan digunakan dalam pengolahan.
- b. Max. Interaksi, digunakan untuk memasukkan jumlah maksimum interaksi yang akan digunakan dalam pengolahan. Ada 3 pilihan yaitu 1,2,3.
- c. Min. Observasi, digunakan untuk memasukkan jumlah minimum observasi yang akan digunakan dalam pengolahan. Ada beberapa pilihan yaitu default, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. Ketika pengguna memilih default, maka program akan menghitung otomatis berdasarkan banyaknya observasi data dan banyaknya prediktor yang digunakan (Friedman, 1991).

3. Output MARS BiresponKontinu (MARS-BK)

Tombol “Run MARS-BK & Simpan Hasil”, digunakan untuk melakukan proses pemodelan MARS birespon kontinu dan menyimpan hasil output dan modelnya dalam lokasi direktori file yang pengguna/user inginkan. Dari tombol ini diperoleh nilai $RMSE$, R^2 , dan nama direktori file tempat pengguna/user menyimpan hasil output dan model program.

4. Output Regresi Parametrik

Tombol “Run Regresi Parametrik”, digunakan untuk melakukan proses pemodelan data dengan regresi parametrik baik regresi linier, kuadratik dan kubik. Dari tombol ini diperoleh nilai $RMSE$ dan R^2 -adj masing-masing untuk regresi linier, kuadratik dan kubik. Dari output yang dihasilkan, dapat diketahui perbandingan ketepatan estimatornya dengan model MARS birespon kontinu.

Halaman ini sengaja dikosongkan

DAFTAR RIWAYAT HIDUP

A. Data Pribadi



Nama : Ayub Parlin Ampulembang
NRP : 1312301702
Tempat / Tgl. lahir : Makassar / 25 Januari 1977
Alamat : Jl. Baji Gau I No. 32 Makassar
Instansi : Badan Pusat Statistik
Provinsi Sulawesi Selatan
NIP : 197701251999011001
Email : ayub.parlin@bps.go.id
Nomor HP : 085257274872
Tgl. Masuk S3 : 3 September 2012
Tgl. Kualifikasi : 23 Juni 2014

B. Riwayat Pendidikan

- S2 Perencanaan Kependudukan, Universitas Hasanuddin (UNHAS), Makassar, Lulus Tahun 2005
- S1 Statistik, Universitas Terbuka (UT), Jakarta, Lulus Tahun 2000
- D3 Statistik, Akademi Ilmu Statistik (AIS), Jakarta, Lulus Tahun 1998
- SMAN 2, Makassar, Lulus Tahun 1995
- SMPN 3, Makassar, Lulus Tahun 1992
- SDN Kompleks Sambung Jawa, Makassar, Lulus Tahun 1989

C. Riwayat Publikasi

Berikut ini publikasi yang telah dilakukan dalam bentuk jurnal dan seminar selama masa studi:

Jurnal Internasional :

- Ampulembang, A.P., Otok, B.W., Rumiati, A.T., Budiasih (2015). Biresponses Nonparametric Regression Model Using MARS and Its Properties. *Applied Mathematical Sciences*, Vol. 9, 2015, no. 29, 1417-1427, ISSN: 1312-885X (Print), ISSN: 1314-7552 (Online), doi: 10.12988/ams, <http://dx.doi.org/10.12988/ams.2015.5127>.
- Ampulembang, A.P., Otok, B.W., Rumiati, A.T., Budiasih (2015). Modeling of Welfare Indicators in Java Island using Biresponse MARS. *International Journal of Applied Mathematics and Statistics (IJAMAS)*, Volume 54, Issue 2, 2016. ISSN 0973-1377 (Print), ISSN 0973-7545 (Online), <http://www.ceser.in/ceserp/index.php/ijamas/article/view/3973>.

Seminar Internasional :

- Ampulembang, A.P., Otok, B.W., Rumiati, A.T., Budiasih (2014). Comparison of MARS and Truncated Spline Approach for Modeling Human Development Index (HDI) in Indonesia. *Proceedings of International Conference on Statistics and Mathematics (ICSM 2014)*. Diselenggarakan oleh Jurusan Statistika FMIPA Institut Teknologi Sepuluh Nopember (ITS) pada tanggal 27-28 November 2014 di Hotel Majapahit, Surabaya, Indonesia. ISSN : 2460 – 0237.
- Ampulembang, A.P., Otok, B.W., Rumiati, A.T., Budiasih (2015). Simulation Study for Biresponses Nonparametric Regression Model using MARS. *Proceedings of The 7th SEAMS-UGM International Conference on Mathematics and Its Applications 2015*, AIP Conf. Proc. 1707, 080003(2016); doi: 10.1063/1.4940860. (<http://dx.doi.org/10.1063/1.4940860>). Diselenggarakan oleh Jurusan Matematika FMIPA Universitas Gadjah Mada (UGM) pada tanggal 18-21 Agustus 2015 di Ruang Seminar Magister Management (MM) UGM Yogyakarta.
- Otok, B. W., Ampulembang, A. P. (2016). Anova Decomposition and Importance Variabel Process in Multivariate Adaptive Regression Splines Model. *International Conference on Theoretical and Applied Statistics (ICTAS 2016)*. Diselenggarakan oleh Jurusan Statistika FMIPA Institut Teknologi Sepuluh Nopember (ITS) pada tanggal 19-20 Oktober 2016 di Hotel Majapahit, Surabaya, Indonesia.

Seminar Nasional :

- Ayub Parlin Ampulembang, Bambang Widjanarko Otok, Agnes Tuti Rumiati & Budiasih (2015). Pemilihan Basis Fungsi Optimal pada Estimator MARS dalam Regresi Nonparametrik Birespon. *Prosiding Seminar Nasional Matematika dan Pendidikan Matematika (SENDIKA)*, Volume 1/ No. 1/2015. Diselenggarakan oleh Program Studi Pendidikan Matematika FKIP Universitas Muhammadiyah Purworejo (UMP) pada tanggal 9 Mei 2015 di Ruang Seminar UM Purworejo, ISSN : 2459-962X.